Stat 100a: Introduction to Probability.

Outline for the day:

1. Collect hw.
2. Bayes's rule.
3. Pot odds.
4. SD and variance.
5. Bernoulli random variables.
6. Binomial random variables.
7. Geometric random variables.
8. Poisson random variables.
9. Continuous random variables, pdf, expected value, and variance.
10. Moment generating functions.
11. Exponential distribution.
12. Independent random variables.

Read through chapter 6.4!

♣ ♥ ◆ ♠

1. Turn in HW.

2. Bayes's rule, p49-52.

Suppose that B_1, B_2 , B_n are disjoint events and that exactly one of them must occur.

Suppose you want $P(B_1 | A)$, but you only know $P(A | B_1)$, $P(A | B_2)$, etc.,

and you also know $P(B_1)$, $P(B_2)$, …, $P(B_n)$.

Bayes' Rule: If $B_1, \ldots, B_n$ are disjoint events with $P(B_1 \text{ or } \ldots \text{ or } B_n) = 1$, then

$P(B_i | A) = P(A | B_i) * P(B_i) \div [\sum P(A | B_j)P(B_j)]$.

Why? Recall: $P(X | Y) = P(X \& Y) \div P(Y)$. So $P(X \& Y) = P(X | Y) * P(Y)$.

$$P(B_1 | A) = P(A \& B_1) \div P(A)$$

$$= P(A \& B_1) \div [\, P(A \& B_1) + P(A \& B_2) + \ldots + P(A \& B_n) \,]$$

$$= P(A | B_1) * P(B_1) \div [\, P(A | B_1)P(B_1) + P(A | B_2)P(B_2) + \ldots + P(A | B_n)P(B_n) \,].$$

Bayes's rule, continued.

Bayes's rule: If $B_1, \ldots, B_n$ are disjoint events with $P(B_1$ or … or $B_n) = 1$, then

$P(B_i \mid A) = P(A \mid B_i) * P(B_i) \div [\sum P(A \mid B_j)P(B_j)]$.

See example 3.4.1, p50. If a test is 95% accurate and 1% of the pop. has a condition, then given a random person from the population,

P(she has the condition | she tests positive)

= P(cond | +)

= P(+ | cond) P(cond) ÷ [P(+ | cond) P(cond) + P(+ | no cond) P(no cond)]

= 95% x 1% ÷ [95% x 1% + 5% x 99%]

~ 16.1%.

Tests for rare conditions must be extremely accurate.

Bayes' rule example.

Suppose P(your opponent has the nuts) = 1%, and P(opponent has a weak hand) = 10%. Your opponent makes a huge bet. Suppose she'd only do that with the nuts or a weak hand, and that P(huge bet | nuts) = 100%, and P(huge bet | weak hand) = 30%.

What is P(nuts | huge bet)?

P(nuts | huge bet) =

$$\frac{P(\text{huge bet} \mid \text{nuts}) * P(\text{nuts})}{P(\text{huge bet} \mid \text{nuts})\, P(\text{nuts}) + P(\text{huge bet} \mid \text{horrible hand})\, P(\text{horrible hand})}$$

$$= \frac{100\% * 1\%}{100\% * 1\% + 30\% * 10\%}$$

= **25%.**

3. POT ODDS CALCULATIONS.

Suppose someone bets (or raises) you, going all-in. What should your chances of winning be in order for you to correctly call?

Let B = the amount bet to you, i.e. the additional amount you'd need to put in if you want to call. So, if you bet 100 & your opponent with 800 left went all-in, B = 700.

Let POT = the amount in the pot right now (including your opponent's bet).

Let p = your probability of winning the hand if you call. So prob. of losing = 1-p.

Let CHIPS = the number of chips you have right now.

If you call, then E[your chips at end] = (CHIPS - B)(1-p) + (CHIPS + POT)(p)

= CHIPS(1-p+p) - B(1-p) + POT(p) = CHIPS - B + Bp + POTp

If you fold, then E[your chips at end] = CHIPS.

You want your expected number of chips to be maximized, so it's worth calling if

-B + Bp + POTp > 0, i.e. if **p > B / (B+POT).**

3) Pot odds and expected value, continued.

From previous slide, to call an all-in, need P(win) > B ÷ (B+pot).
Expressed as an *odds ratio*, this is sometimes referred to as *pot odds* or *express odds*.

If the bet is not all-in & another betting round is still to come, need
P(win) > wager ÷ (wager + winnings),
where winnings = pot + amount you'll win on later betting rounds,
wager = total amount you will wager including the current round & later rounds, assuming no folding.

The terms *Implied-odds* / *Reverse-implied-odds* describe the cases where *winnings > pot* or where *wager > B*, respectively. See p66.

Example: 2006 World Series of Poker (WSOP). ♣ ♥ ♦ ♠

Blinds: 200,000/400,000, + 50,000 ante.

Jamie Gold (4♣ 3♣): 60 million chips. Calls.

Paul Wasicka (8♠ 7♠): 18 million chips. Calls.

Michael Binger (A♥ 10♥): 11 million chips. Raises to $1,500,000.

Gold & Wasicka call. (pot = 4,650,000) Flop: 6♠ 10♣ 5♠.

•Wasicka checks, Binger bets $3,500,000. (pot = 8,150,000)

•Gold moves all-in for 16,450,000. (pot = 24,600,000)

•Wasicka folds. Q: Based on expected value, should he have called?

If Binger will fold, then Wasicka's chances to beat Gold must be at least 16,450,000 / (24,600,000 + 16,450,000) = 40.1%.

If Binger calls, it's a bit complicated, but basically Wasicka's chances must be at least 16,450,000 / (24,600,000 + 16,450,000 + 5,950,000) = 35.0%.

4. Variance and SD.

Expected Value: $E(X) = \mu = \sum k\ P(X=k)$.

Variance: $V(X) = \sigma^2 = E[(X- \mu)^2]$. Turns out this $= E(X^2) - \mu^2$.

Standard deviation $= \sigma = \sqrt{V(X)}$. Indicates how far an observation would *typically* deviate from μ.

Examples:

<u>Game 1.</u> Say X = \$4 if red card, X = \$-5 if black.

E(X) = (\$4)(0.5) + (\$-5)(0.5) = -\$0.50.

$E(X^2)$ = ($\$4^2$)(0.5) + ($\$-5^2$)(0.5) = (\$16)(0.5) + (\$25)(0.5) = \$20.5.

So $\sigma^2 = E(X^2) - \mu^2$ = \$20.5 - $\$-0.50^2$ = \$20.25. **σ = \$4.50**.

<u>Game 2.</u> Say X = \$1 if red card, X = \$-2 if black.

E(X) = (\$1)(0.5) + (\$-2)(0.5) = -\$0.50.

$E(X^2)$ = ($\$1^2$)(0.5) + ($\$-2^2$)(0.5) = (\$1)(0.5) + (\$4)(0.5) = \$2.50.

So $\sigma^2 = E(X^2) - \mu^2$ = \$2.50 - $\$-0.50^2$ = \$2.25. **σ = \$1.50**.

5. Bernoulli Random Variables, ch. 5.1.

If X = 1 with probability p, and X = 0 otherwise, then X = *Bernoulli (p)*.

Probability mass function (pmf):

$P(X = 1) = p$

$P(X = 0) = q$, where $p+q = 100\%$.

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

For example, suppose X = 1 if you have a pocket pair next hand; X = 0 if not.

$p = 5.88\%$. So, $q = 94.12\%$.

[Two ways to figure out p:

(a) Out of choose(52,2) combinations for your two cards, 13 * choose(4,2) are pairs.

13 * choose(4,2) / choose(52,2) = 5.88%.

(b) Imagine *ordering* your 2 cards. No matter what your 1st card is, there are 51 equally likely choices for your 2nd card, and 3 of them give you a pocket pair. 3/51 = 5.88%.]

$\mu = E(X) = .0588$. SD = $\sigma = \sqrt{(.0588 * 0.9412)} = 0.235$.

6. 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).

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?</u> X = Binomial (100, 5.88%).

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

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

E(X) = np = 100 * 0.0588 = **5.88**. σ = √(100 * 0.0588 * 0.9412) = **2.35**.

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

7. 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.

8. Poisson random variables, ch 5.5.

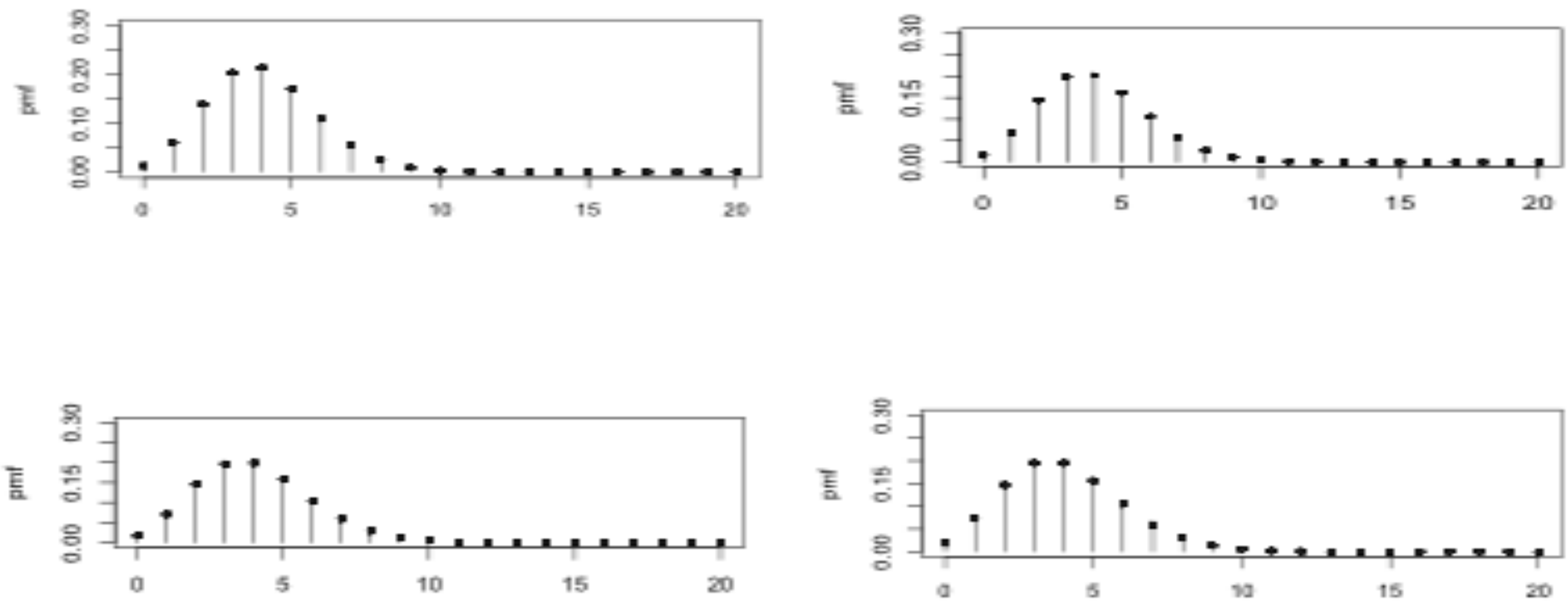
Player 1 plays in a very slow game, 4 hands an hour, and she decides to do a big bluff whenever the second hand on her watch, at the start of the deal, is in some predetermined 10 second interval.

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 4 second interval.

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 2 second interval. Each of the three players will thus average one bluff every hour and a half.

Let X_1, X_2, and X_3 denote the number of big bluffs attempted in a given 6 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, 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) $P(X = 5 \mid X > 1) = P(X = 5 \text{ and } X > 1) \div P(X > 1) = P(X = 5) \div P(X>1) = [e^{-1.4}\ 1.4^5/5!] \div [1 - e^{-1.4}\ 1.4^0/0! - e^{-1.4}\, 1.4^1/1!] \sim 2.71\%$.

9. Continuous random variables and their densities, p103-107.

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

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

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

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

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

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

10. 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.

1^{st} 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.

$\phi_X(t) = E(e^{tX})$ 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?

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

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

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

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

$= 0.72 + 0.28e^t$, 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.

11. Exponential distribution, ch 6.4.

Useful for modeling waiting times til something happens (like the geometric).

pdf of an exponential random variable is $f(y) = \lambda\ exp(- \lambda\ y)$, for $y \geq 0$, and $f(y) = 0$ otherwise.
If X is exponential with parameter λ, then $E(X) = SD(X) = 1/\lambda$

If the total numbers of events in any disjoint time spans are independent, then these totals are Poisson random variables. If in addition the events are occurring at a constant rate λ, then the times between events, or *interevent times*, are exponential random variables with mean $1/\lambda$.

Example. Suppose you play 20 hands an hour, with each hand lasting exactly 3 minutes, and let X be the time in hours until the end of the first hand in which you are dealt pocket aces. Use the exponential distribution to approximate $P(X \leq 2)$ and compare with the exact solution using the geometric distribution.

Answer. Each hand takes 1/20 hours, and the probability of being dealt pocket aces on a particular hand is 1/221, so the rate λ = 1 in 221 hands = 1/(221/20) hours ~ 0.0905 per hour.
Using the exponential model, $P(X \le 2 \text{ hours}) = 1 - \exp(-2\lambda) \sim 16.556\%$.
This is an approximation, however, since by assumption X is not continuous but must be an integer multiple of 3 minutes.
Let Y = the number of hands you play until you are dealt pocket aces. Using the geometric distribution, $P(X \le 2 \text{ hours}) = P(Y \le 40 \text{ hands})$ $= 1 - (220/221)^{40} \sim 16.590\%$.

The survivor function for an exponential random variable is particularly simple: $P(X > c) = \int_c^\infty f(y)dy = \int_c^\infty \lambda \exp(-\lambda y)dy = -\exp(-\lambda y)]_c^\infty = \exp(-\lambda c)$.

Like geometric random variables, exponential random variables have the *memorylessness* property: if X is exponential, then for any non-negative values a and b, $P(X > a+b \mid X > a) = P(X > b)$.
Thus, with an exponential (or geometric) random variable, if after a certain time you still have not observed the event you are waiting for, then the distribution of the *future,* additional waiting time until you observe the event is the same as the distribution of the *unconditional* time to observe the event to begin with.

12. Independent random variables.

If X and Y are independent random variables, then

$E[f(X)\, g(Y)] = E[f(X)]\, E[g(Y)]$, for any functions f and g.

See Exercise 7.12. This is useful for problem 5.4.