Announcements

- the course schedule
- we will switch to using the standard UCLA Course Web site for this class, to post class materials. We'll put a link on the old class website to the UCLA Course Web site for this class.
- Want more lectures for this class? My "traditional" lectures for this class have been posted as videos online. We'll create a page with links to all the lectures.

Lessons from the Exercises

Conditional Probability (Disease Test)

- Many people didn't consider the *direction* of the conditional probability, even though the question's phrasing and answers encouraged you to do that. (The question gave you $p(O|H)$ but asked you about $p(H|O)$). Implies they didn't realize that any conditional probability has two possible directions.
- Those who did consider this fact immediately saw that the numbers implied a very poor reliability for positive test results (<25%, due to false positives). No calculation was required to see this!
- Some people chased red herrings like "does *reliable* mean 95%? 97%?"

Key Lessons

- Etch into your minds: *Which variable is hidden? Which varible is observed? Which direction of conditional probability am I being asked for?*
- Etch into your minds: if a (hidden) state is rare, be very worried about the false positive rate (no matter how good the test is)!!

Joint Probability

- common error: confusing "joint probability" with "conditional probability". Watch out! $p(O,H) \neq p(O|H)$. Some people wrote $p(O|H)$, others wrote $p(H|O)$. Not what the question asked!

	T^-	T^+	total
D^+	1	9	10
D^-	960	30	990
total	961	39	1000

Table: A diagnostic disease test: 1000 patients were given a diagnostic test that gives either a positive () or negative () result, and independently assessed for whether they have the disease () or not () by rigorous clinical criteria.

Monty Hall Observational Independence

Mistakes we made:

- **B**. sweeping generalization: "obs does give information about where the prize is". Yes, but that's not what the question asked. It asked about whether *door A contains the prize*, i.e. h is $\delta = A$. And in fact $p(\delta = A|obs) = p(\delta = A)$.
  Evidently just not thinking about the independence argument that this question stated.
- a few people asserted $p(\delta = A|obs) = 1/2$ (because only two doors are left). Still laboring under the misconception that all (closed) doors have equal probability after observing B^-!

Mistakes we made:

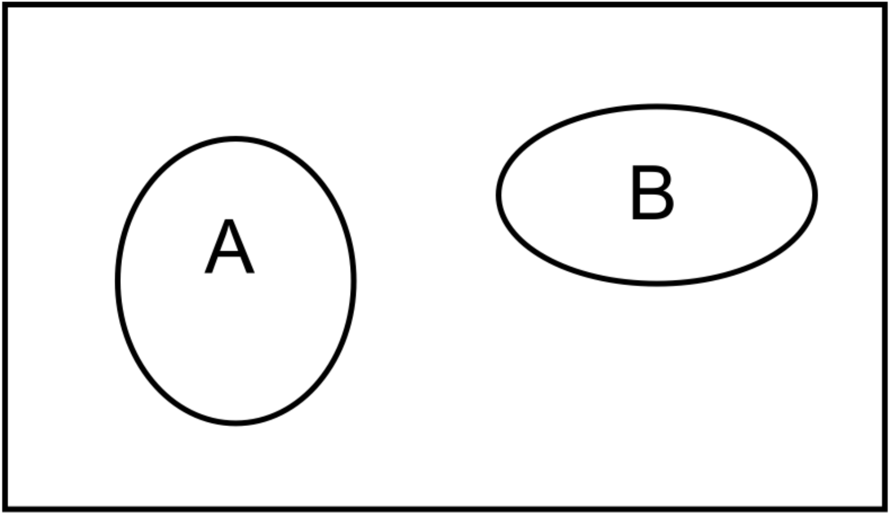
- Beware of summing over the condition of a conditional probability!! This is a common error.
- Many people realized 2 and 3 were equal, because $p(A,B) = p(A|B)p(B)$, but then didn't see the connection to 5, 6 ($p(A)$). This implies they didn't realize that summing over B eliminates it from the probability expression.
  This is a *really* important tool. It allows you to properly handle a hidden variable that links what you know (obs) and what you want to know (some other hidden variable). Just sum over all possible values of this "nuisance variable" and eliminate it from the equation!!
- Some people assumed A,B were independent, contrary to what the question asked.

Using Posteriors as Priors

Mistakes we made:

- **A**. intuitive error: “looks analogous to Bayes law, so it seems reasonable”. Not even considering the question of independence.
- **B**. sweeping generalization: $p(A, B) \neq p(A)p(B)$ treated as meaning they can never be equal. But if A, B independent, they are equal.
- variable vs. event confusion: “X1 has to intersect with x2 but they do not”. X_1, X_2 are two random *variables*. Talking about “intersection” only makes sense for two *states*. Don’t mix these up!!

Event Independence?



Are events A, B statistically independent? Justify your answer mathematically.

Drawing Multiple Variables At Once

- When we draw a point from a joint distribution $p(A, B, C, D)$, each draw is a tuple of values a, b, c, d.
- Think of this as "throwing a dart at a dartboard" representing our total set. For each variable that we've defined on that set, that point lies in exactly one disjoint slice that corresponds to one value of that variable. So that one point gives us a value of *each* of the variables we've defined on that set.

Variable Independence

For a pair of random variables X, Y, you are told that $p(Y|X) = p(Y)$. Can you automatically conclude from this that $p(X|Y) = p(X)$?

1. Yes, always
2. No, never
3. Sometimes

Independence as "Reduced Information"

- say X, Y each have n states.
- the joint probability table $p(X, Y)$ has $O(n^2)$ degrees of freedom.
- the independent product $p(X)p(Y)$ has only $O(2n)$ degrees of freedom.
- there are $O(n^2)$ conditional probabilities $p(X|Y)$.

Independence is about the entire joint distribution

Independence means "zero information" in the sense that X gives *absolutely no information* about Y.

- Even the slightest bias in just a part of the joint distribution would violate that definition, i.e. *not* strictly independent.
- Think of this like *correlation*, which you measure by plotting the entire dataset of X, Y values. If they show some correlation, they each give information about each other. Absolutely no correlation equals independence.

Rules of the Road

- you can always write a joint probability in terms of the chain rule, in any order you like...

$$p(X_1, X_2, ...X_n) = p(X_1)p(X_2|X_1)...p(X_n|X_1, X_2, ...X_{n-1})$$

- you can always derive any marginal (unconditional, e.g $p(X_1, X_2)$) or conditional probability (e.g. $p(X_n|X_1)$) you like...
- but you cannot assume that some particular product of marginals and conditionals will correctly give you the joint probability, e.g. $p(X_1)p(X_2)...p(X_n)$. Only true if that *specific independence assertion* is actually true, across the whole joint probability function!

Posterior Effect

You are analyzing a genetic test that tests whether a woman is a carrier (or not) of a rare genetic disease (found in less than 1% of the population). Say you observe a test result that is four times as likely if she is a carrier than if she is not. How will that alter your estimate of the probability that she is a carrier? (State approximately what you expect the posterior / prior ratio to be, or say “insufficient data”).