Conditional Probabilities

A study of treatments for depression

	Treatment Group				
Response	*Imipramine*	*Lithium*	*Comb.*	*Placebo*	**Total**
Relapse	18	13	22	24	77
No relapse	22	25	16	10	73
Total	40	38	38	34	150

- What's the probability that a patient who received the placebo had a relapse?
- Let B be the event that a patient had a relapse, and A the event that a patient received the placebo, find the probability of the event B, given that we know event A has occurred.
- How does this quantity relate to the efficacy of the treatment(s)?

The conditional probability of event A given event B

- $P(A|B) = \dfrac{P(AB)}{P(B)}$, where $P(B) > 0$.

  We consider the "reduced sample space" to be the set B, and determine the fraction of that sample space that is occupied by A.
- Example: One die is rolled twice in a rox. Let B be the event that the first roll is a 3, and let A be the event that the second roll is a 3.

  Compare $P(A), P(AB), P(A|B)$.

The Multiplication Rule: using conditional probabilities to find the probability of intersections.

Suppose an urn contains r red balls and b blue balls, and two balls are drawn without replacement. What is the probability of getting first a blue ball and then a red ball?

- The definition of conditional probability gives:

$$P(AB) = P(B)P(A|B)$$

  Interpret in words, where A is the event that the second ball is red, and B is the event that the first ball is blue.
- Using the formula, $P(AB) =$ ________________________.
- This principle can be generalized:

$$P(A_1A_2\cdots A_n) = P(A_1)P(A_2|A_1)P(A_3|A_1A_2)\ldots P(A_n|A_1A_2\cdots A_{n-1})$$

  where $P(A_1A_2\cdots A_{n-1}) > 0$.
- *Proof . . .*

Ex: What's the probability of winning at craps?

- **Rules:** Two dice are rolled and summed to S_1.
  • If $S_1 = 7$ or 11, then player wins.
  • If $S_1 = 2, 3, 12$, then player loses.
  • Otherwise, dice are rolled repeatedly and summed to get $S_2, S_3, \dots$. Player wins if $S_i = S_1$ *before* $S_i = 7$.
- Find:
  $\pi_0 =$ probability of winning on first roll.
  $\pi_k =$ probability of winning when first roll is k, for $k = 4, 5, 6, 8, 9, 10$.

Definition of Independent Sets

- A and B are if and only if $P(AB) = P(A)P(B)$.

  If $P(B) > 0$, this can be written:
  A and B independent $\Leftrightarrow P(A|B) = P(A)$.
  Interpret in words.
  If two sets are not independent, they are **dependent**.
- The sets $A_1, A_2, \dots A_n$ are independent if and only if:

$$P(A_1 A_2 \cdots A_n) = P(A_1)P(A_2)\dots P(A_n)$$

- Ex.: An infinite sequence of independent trials is to be performed. Each trial has a probability p of success. What is the probability that:
  a) at least 1 success occurs in the first n trials;
  b) exactly k successes occur in the first n trials;
  c) all trials result in success.
  See also Examples 2.2.4 and 2.2.7 in text.

Are the following sets, E and F, independent? *from Ross*

- E = event a businesswoman has blue eyes, F = even her secretary has blue eyes.
- E = event a professor owns a car, F = event the professor is listed in the phone book
- E = event a man is at least 6 ft. tall, F = event he weighs over 200 pounds
- E = event a dog lives in the United States and F is the event that the dog lives in the western hemisphere.
- E = event it will rain tomorrow, F = event it will rain the day after tomorrow

Example: Coupon Collecting

Suppose there are r different coupons, independently assigned to packages (e.g. prizes in CrackerJacks, Coke bottle tops, baseball cards in bubble gum). What is the probability that a complete set (of r coupons) will be obtained from n randomly chosen packages?

- Let A_i denote the event that the ith coupon is missing from all n packages.
- Write the probability of obtaining a complete set in terms of these A_is.
- Find $P(A_i)$ and $P(A_iA_j)$.
- Answer the problem.

Law of Total Probability

- **Definition:** A *partition* of a sample space, S, is a collection of *disjoint* sets, $\{B_i\}_{i=1}^n$ whose union is all of S.
- **Law of total probability** Suppose $\{B_i\}_{i=1}^n$ is a partition of S such that $P(B_i) > 0$ for all i. Then, for every $A \subset S$:

$$P(A) = \sum_{i=1}^{n} P(B_i)P(A|B_i)$$

- **Conditional version:**

$$P(A|C) = \sum_{i=1}^{n} P(B_i|C)P(A|B_iC)$$

- In the clinical trial for treatments for depression, suppose that, in a particular treatment group, Prob(no relapse) = p, but we don't know p. Partition the sample space into sets B_i such that $\frac{i-1}{10} = p$ in $B_i, i = 1, 2, \ldots 11$. For a particular person from a particular group, e.g. the iprimamine group, let $E_1 =$ event this person has no relapse. By assumption, $P(E_1|B_i) = \frac{i-1}{10}$. If we have *no* information about p, then all

Bayes Theorem

Let $\{B_i\}_{i=1}^n$ be a partition of S such that $P(B_i) > 0$ for all i, and let A be a set such that $P(A) > 0$. Then:

$$P(B_i|A) = \frac{P(B_i)P(A|B_i)}{\sum_{j=1}^{k} P(B_j)P(A|B_j)}$$

Proof: ...

Thomas Bayes (1702-1761)

In the clinical trial example, we observe that the first patient did *not* have a relapse. Recall that we assumed that, if all B_i were equally likely, $P(E_1) = .5$. How does the data change our belief? I.e., how does it change the probabilities of the B_is?

Laplace's rule of succession

 1749-1827

There are $k+1$ coins in a box, where the ith coin has probability i/k of coming up heads when tossed, $i = 0, 1, \ldots k$. A coin is selected and repeatedly flipped. If it comes up heads for the first n tosses, what is the probability that the next toss will also be a head?