Math 1070-2: Spring 2008

Lecture 5

Davar Khoshnevisan

Department of Mathematics
University of Utah
http://www.math.utah.edu/~davar

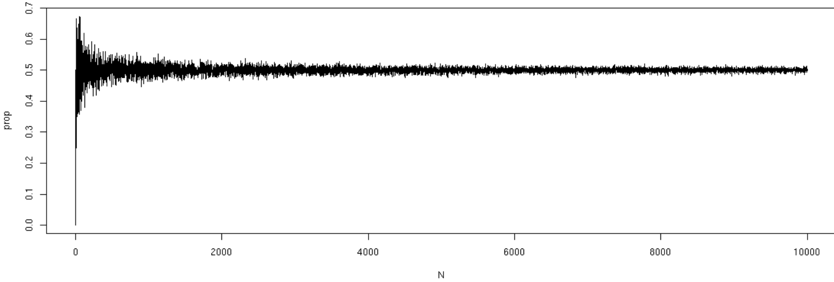
February 6, 2008

Randomness

- Simulations of a fair coin (1 = heads; 0 = tails)
  - 0 0 0 0 1 1 0 0 1 1 [0.4=40%]
  - 1 1 1 0 1 1 1 0 0 0 [0.6=60%]
  - 0 1 0 0 0 0 1 1 0 0 [0.3=30%]
  - 0 0 0 0 0 1 1 0 1 0 [0.3=30%]
  - 1 0 0 0 0 0 1 0 0 0 [0.3=30%]
  - 1 0 0 0 1 0 0 0 1 1 [0.4=40%]
  - 1 0 1 0 0 0 0 0 1 1 [0.4=40%]
  - 0 1 1 1 1 0 1 1 1 0 [0.7=70%]
  - 0 0 1 1 0 1 0 1 1 0 [0.5=50%]
- Is this random? Is this a fair coin?

Randomness

- Toss N fair coins; tally the proportion of heads
- Long-run pattern



Probability

- Assignment of likelihood
- Usually has a long-run interpretation
  [Law of large numbers; J. Bernoulli, 1689]
  - Probability of heads in a toss of a fair coin
  - Probability of rolling two dots in a roll of a fair die
  - Probability of rain tomorrow (??)
  - Probability that candidate X wins the next election (??????)

Equally-likely outcomes

- Example to have in mind: Toss a fair coin 3 times. What is the probability of getting 2 heads?
- To compute probabilities of equally likely events:
  - Produce [usually in your head] a "sample space" [this is a list of all possible outcomes of the experiment]
    - Here, we could choose:
      TTT TTH THT THH
      HHH HHT HTH HTT
  - Each element of the sample space is equally likely [there are other possibilities; this is good enough for us]
  - Prob=# ways to get the outcome we want /total # possible outcomes
    - Here,

$$P(\text{two heads}) = \frac{3}{8} = 0.375 = 37.5\%$$

Equally-likely outcomes

- Had the following sample space:
  TTT TTH THT THH
  HHH HHT HTH HTT $P(\text{two heads}) = \frac{3}{8}$
- What if we wrote the sample space unordered?
  TTT TTH THH HHH $P(\text{two heads}) = \frac{1}{4}$
- These can't *both* be right. What is going on?

Independence

- Events A and B are **independent** if:
  - $P(B)$ is the same as the probability of B if you were told A
  - Consider our old sample space
    TTT TTH THT THH HHH HHT HTH HTT
  - Let A be the event first coin-toss is heads
  - Let B be the event third coin-toss is tails
  - Are A and B independent?
  - The probability of B is $P(B) = \frac{4}{8} = \frac{1}{2}$
  - If you knew A then the sample space is reduced to
    HHH HHT HTH HTT
  - The [conditional] probability of B [given A] in this case is
    $P(B|A) = \frac{2}{4} = \frac{1}{2}$
  - Yes! A and B are independent
- Two draws from a deck of cards. Are the draws independent?

Independence

- A precise mathematical definition:
  - *A* and *B* are independent if

$$P(A\&B) = P(A)P(B)$$

- A more-or-less honest verbal definition:
  - *A* and *B* are independent if these outcomes don't affect each other [statistically speaking, whatever that means]

Independence and sampling

- Sampling with replacement $\longrightarrow$ independent draws
- Sampling without replacement $\longrightarrow$ draws that are *not* independent
- If the population is large then both sampling methods $\longrightarrow$ independent draws

Random variables and distributions

- A **random variable** is the as-yet unseen outcome of a random experiment
- Its **distribution** is a list of two quantities:
  - Possible values
  - Versus probabilities
- There are also “continuous” random variables [later]

Example: A die

possible value	probability
1	$1/6 \approx 0.17$
2	$1/6 \approx 0.17$
3	$1/6 \approx 0.17$
4	$1/6 \approx 0.17$
5	$1/6 \approx 0.17$
6	$1/6 \approx 0.17$
	sum =1

Example: # of homeruns in a game for the Red Sox

("based on 2004 data")

possible value	probability
0	0.23
1	0.38
2	0.22
3	0.13
4	0.03
5	0.01
6 or more	0.00
	sum =1

- What does this mean? How is it computed?

The mean (expectation) of a probability distribution

- A box has 3 ones and 2 threes. The average value in the box is

$$\frac{1+1+1+3+3}{5} = \frac{9}{5} = \left(1 \times \underbrace{\frac{3}{5}}_{P(1)}\right) + \left(3 \times \underbrace{\frac{2}{5}}_{P(3)}\right)$$

- $\frac{9}{5} = 1.8$ is our best guess for the outcome of this draw [before it happens]
- General formula:

$$\mu = \sum_{x=\text{possible values}} xP(x)$$

Example: A die

possible value	probability
1	$1/6 \approx 0.17$
2	$1/6 \approx 0.17$
3	$1/6 \approx 0.17$
4	$1/6 \approx 0.17$
5	$1/6 \approx 0.17$
6	$1/6 \approx 0.17$
	sum =1

$$\mu = \left(1 \times \frac{1}{6}\right) + \left(2 \times \frac{1}{6}\right) + \cdots + \left(6 \times \frac{1}{6}\right) = 3.5$$

- Does this make a good guess? For what? And how?

Example: # of homeruns in a game for the Red Sox

("based on 2004 data")

possible value	probability
0	0.23
1	0.38
2	0.22
3	0.13
4	0.03
5	0.01
6 or more	0.00
	sum =1

- $$\mu = (0 \times 0.23) + (1 \times 0.38) + (2 \times 0.22) + \cdots + (5 \times 0.01) \approx 1.38$$

Law of large numbers (again)

- General fact: Take a large independent sample from a population, and consider a random variable that one would obtain in this way [e.g., weight]
- Then the sample average [e.g., sample weight] is [with high probab.] close to the mean of the probab. distribution of the random variable[e.g., true average weight of the population]
- As the sample size $\to \infty$ this approximation gets better, with increasingly improved probabilities

Expectations and taking bets

- We all take bets in different settings. Expectations show us how to do this well.
- Lottery: Costs \$1; win \$0 with probab. $\frac{99{,}999}{100{,}000}$; win \$10,000 with probab. $\frac{1}{100{,}000}$

$$\text{expected win} = \left(0 \times \frac{99999}{100000}\right) + \left(10000 \times \frac{1}{100000}\right) - 1 = -0.9$$

- Lottery 1 $\longrightarrow$ expect to lose 90¢
- Long-run interpretation?