

Probability and Randomness

What is randomness?

A phenomenon or procedure for generating data is random if

- *the outcome is not predictable in advance;*
- *there is a predictable long-term pattern that can be described by the distribution of the outcomes of very many trials.*

A random outcome is the result of a random phenomenon or procedure.

Random experiment—we will not use this term, since “experiment” means something else to us.

Random (regular enough to be modelled) vs Haphazard (too irregular to model effectively).

Theory of Probability

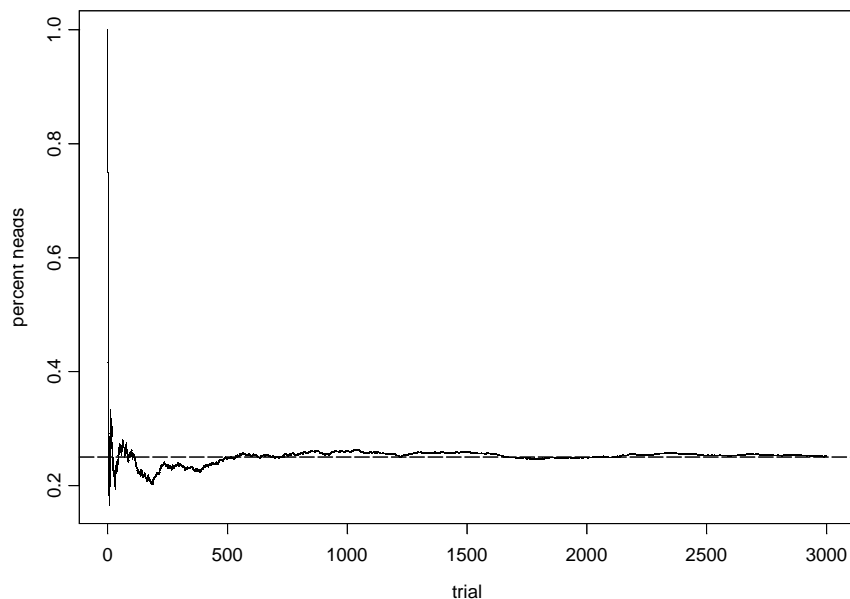
- Astragali and other precursors of dice from ancient times
- Dice, coins etc.
- “Games of chance:” how should a gambler adjust his bet to the odds of winning? What is a “fair bet”?
- Fermat, Pascal (17th cent); later Bernoulli, others.
- Much much later (19th cent? and later) was probability applied to astronomical or human affairs.

The probability of an outcome is a measure of how likely it is for that outcome to occur. Probability must obey two rules:

- (A) *The probability of any outcome of a random phenomenon must be between 0 and 1: $0 \leq p \leq 1$.*
- (B) *The probabilities of all the possible outcomes associated with a particular random phenomenon must add up to exactly 1: $\sum p_i = 1$.*

How do we assign probabilities to outcomes?

- **Subjective** (educated guess) probability: *“I probably won’t pass the test tomorrow.” “I’ve looked at your resume and we very probably will hire you.”*
- **Calculated** (theoretical): *“For an ideal die that is perfectly balanced, the probability of any one face coming up in a toss is $1/6$ (number of faces).” “The data follow the normal distribution so I will use the 68-95-99.7 rule to assign probabilities.”*
- **Empirical** (observed long-run frequency): *coin tossing:*



Calculating Probabilities

- Sample space S : set of all possible outcomes.
- Event A : a (smaller) set of outcomes you are interested in. An event occurs if one of the outcomes in it occurs.

Example: If you toss a six-sided die, the sample space can be written as the list $S = \{1, 2, 3, 4, 5, 6\}$ of possible faces that can come up. One of the events in the sample space is the event $A = \{\text{even face comes up}\}$ which we can also write as $A = \{2, 4, 6\}$. If we perform this procedure and observe that face 4 comes up, we say A has occurred, since one of the outcomes in A , namely 4, occurred.

- The probability of an event A , $P(A)$, is the sum of the probabilities of the outcomes. The probability of S , $P(S)$, is 1.

In the example above, if each face is equally likely to come up (probability $1/6$), then the event A has probability $P(A) = 1/6 + 1/6 + 1/6 = 1/2$.

- The complement of the event A is written A^c ; it is the event consisting of all outcomes not in A (so A^c is the “opposite” of A). The probability of A^c is $P(A^c) = 1 - P(A)$.
- The union of two events A and B is the set of all outcomes in A , or in B , or both, written $A \cup B$. The intersection of A and B is the set of outcomes in **both** A and B , written $A \cap B$.

In a certain introductory statistics class at CMU, the probability of a randomly-sampled student being a freshman was 0.65; the probability of the student being from HSS was 0.70, and the probability of the student being both a freshman and an HSS student was 0.55.

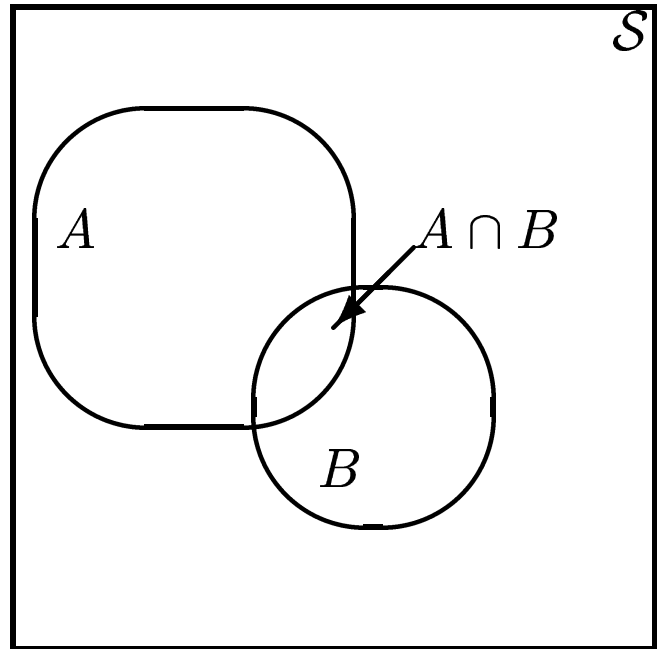
Event A , in words: _____

Event B , in words: _____

$$P(A) = \underline{\hspace{2cm}}$$

$$P(B) = \underline{\hspace{2cm}}$$

$$P(A \cap B) = \underline{\hspace{2cm}}$$



- The general “or” rule is $P(A \cup B) = P(A) + P(B) - P(A \cap B)$. This can easily be seen by equating probabilities with areas in the Venn diagram.

$$P(\text{Fresh}) + P(\text{HSS}) = 0.65 + 0.70 = 1.35 > 1, \text{ double-counts overlap.}$$

$$P(\text{Fresh}) + P(\text{HSS}) - P(\text{Fresh AND HSS}) \\ = 0.65 + 0.70 - 0.55 = 0.80.$$

- The conditional probability of A given B is a way of revising the probability of A once you have the new information that A occurred. It is the fraction of probability in B that is accounted for by A. We write $P(A|B) = \frac{P(A \cap B)}{P(B)}$.

In this class, the probability of being a freshman given that you're in HSS is

$$P[\text{Fresh}|\text{HSS}] = \frac{P[\text{Fresh and HSS}]}{P[\text{HSS}]} = \frac{0.55}{0.70} = 0.79, \text{ approximately.}$$

Joint Probability Table:

		HSS		Total
		Yes	No	
Fresh	Yes	0.55		0.65
	No			
Total		0.70		1.00

Now it is easy to see for example that

- $P[\text{Fresh}|\text{HSS}] = \frac{0.55}{0.70} = 0.79$ is a column percent, and
- $P[\text{HSS}|\text{Fresh}] = \frac{0.55}{0.65} = 0.85$ is a row percent.

- Two events A and B are said to be *independent* if either $P(A|B) = P(A)$ or $P(B|\overline{A}) = P(B)$. In words, the information that B has occurred doesn't change the probability of A , and vice-versa.

$P[\text{Fresh}] = 0.65 \neq 0.79 = P[\text{Fresh}|\text{HSS}]$, so *Freshman status and HSS status are not independent.*

Independence, Need for Regularity

- Since $P[A|B] = P[A \cap B]/P[B]$, it follows that $P[A \cap B] = P[A|B] \cdot P[B]$. (general “and” rule).
- If A and B are independent, $P[A|B] = P[A]$, so $P[A \cap B] = P[A] \cdot P[B]$.

When we think of a sequence of coin flips we usually think of them as *independent*, so the probabilities multiply across flips: $P(HTH) = P(H) \cdot P(T) \cdot P(H)$.

Runs, Human Perception, Hot Hand

- *Which is more likely:*
 - *HTHTHH*
 - *HHHHTT*
- *Hot Hand in Basketball?* Probably not; shots look independent when carefully analyzed:

$P[\text{Shaq makes freethrow}] \approx 0.60$, regardless of what he’s done earlier in the game.

- *Human Perception.* There is probably some survival value in noticing regular patterns (stripes on a tiger sneaking up for attack??).

The Law of Averages

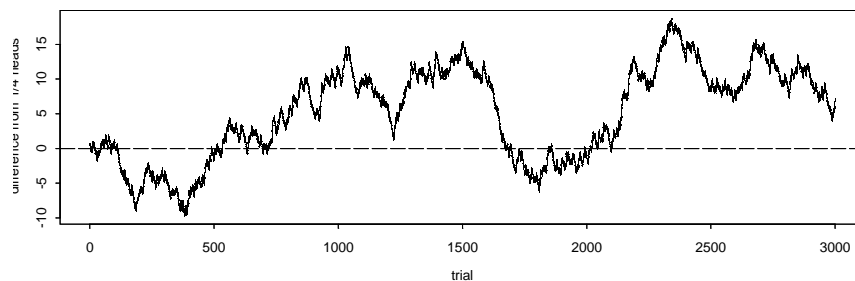
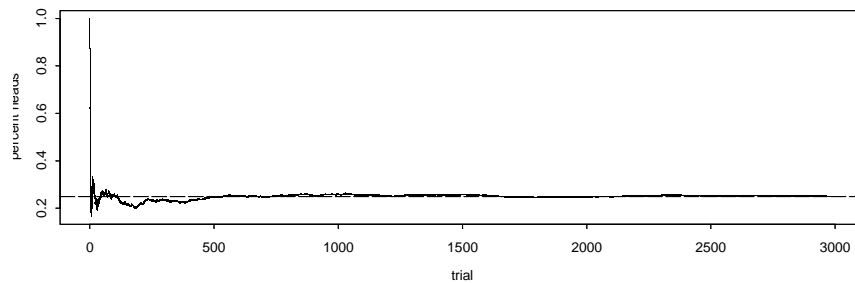
- A baseball player consistently bats .350. The player has failed to get a hit in six straight at-bats; the announcer says “Tony is due for a hit by the law of averages”
- The weatherman notes that for the past several winters temperatures have been mild and there hasn’t been much snow. He observes, “Even though we’re not supposed to use the law of averages, we are due for a cold snowy winter this year!”
- A gambler knows red and black are equally-likely, but he has just seen red come up five times in a row so he bets heavily on black. When asked why, he says “the law of averages!”

If the events are independent, then the probability of seeing a hit, a cold winter, or a black number, is the same this time as it was each previous time. There is no “law of averages.”

This is based on a misunderstanding: that random outcomes achieve regularity by *compensating for past imbalances*. This is wrong.

Instead, random outcomes achieve regularity by *swamping out past imbalances*.

- For any sequence of independent coin tosses, 100 heads followed by 100 tails are just as likely as 100 head and 100 tails all intermingled.
- If we ever start off with 100 heads in a coin-tossing marathon, the number of heads may never be closer to the “expected number” of heads than 100, but the *fraction* of tosses that are heads will approach the “expected fraction”.



Probability and Risk

- Probability can be used to measure risk and make “fair bets”—more on this next time.
- It’s generally difficult for people to use probability this way without some training:
 - We tend to assess risk using heuristics and scripts that may have had some survival value in the past.
Asbestos deaths 15/million; Driving deaths 1500 / million. Yet people tend to prefer risk of driving (“in control”? related to immediate gratification?)
 - Abstract events like asbestos poisoning are difficult to assess the prob. of; even when it can be done, small probabilities are difficult to comprehend.
15/million = 0.000015, vs 1500/million = 0.015
 - People tend to assess probabilities/risks based on individual event built up from “personal” experience, rather than abstract probabilities.
Urban crime is generally rated to be a greater risk of living in the city than it actually is, because it is reported on more frequently in the news.