STAT 340: Introduction to Probability and Mathematical Statistics I

Peter Guttorp
peter@stat.washington.edu
www.stat.washington.edu/peter/340

Class information

Lectures MF
TA section W
Homework due Mondays
Midterm Friday November 5

Grading scheme:
5% class participation
30% homework (drop worst)
30% midterm
35% final (MUST take final)

2

Probability and statistics

Probability is the
science of randomness
–regularity in irregularity
Build probability models to describe observed phenomena

Statistics is a *system for formal inference* based on probability models
Estimate parameters of model
Test hypotheses
Assess uncertainty

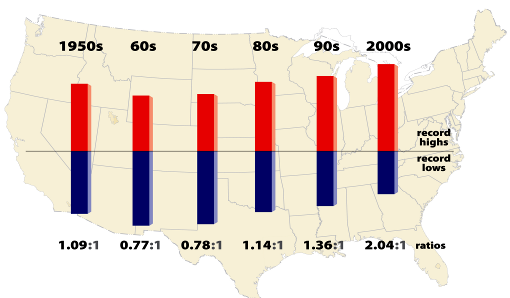
3

Climate change?

IPCC
AR4 WG1 Ch. 11 (2007):
Fewer cold outbreaks; fewer, shorter, less intense cold spells / cold extremes in winter
VL (consistent across model projections)
Northern Europe, South Asia, East Asia
L (consistent with warmer mean temperatures)
Most other regions
Reduced diurnal temperature range
L (consistent across model projections)
Over most continental regions, night temperatures increase faster than day temperatures

4

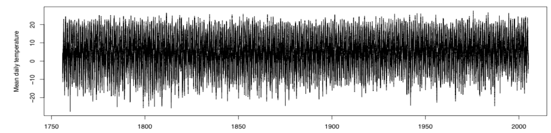
Trends in US record daily temperatures

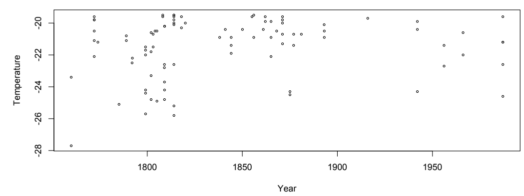


Meehl et al. 2009

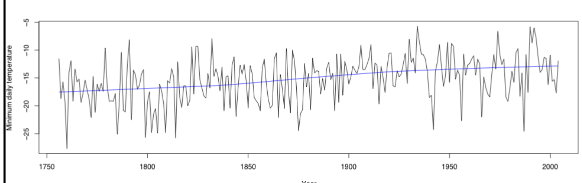
The Stockholm series

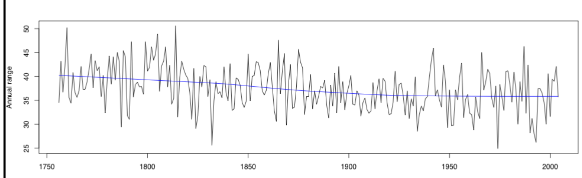
Stockholm daily
1756-2004
Moberg et al. 2002





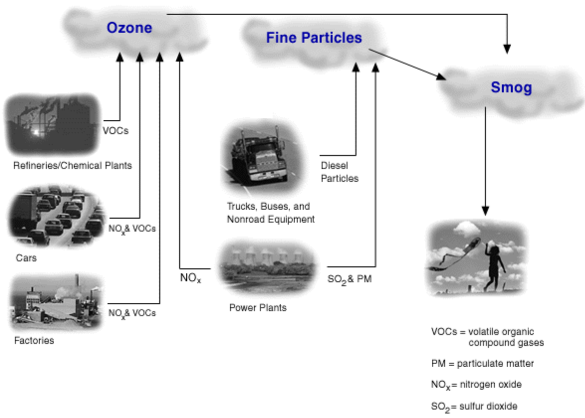
Annual minimum temperature and range





London smog 1952

5 days of pea soup fog
Visibility at times a few feet
12 000 deaths
100 000 made ill



Smog in Beijing

Health effects of ozone

Decreased lung capacity
Irritation of respiratory system
Increased asthma hospital admissions
Children particularly at risk

Ozone regulation

In each region the expected number of daily maximum 1-hr ozone concentrations in excess of 0.12 ppm shall be no higher than one per year

Implementation: A region is in violation if 0.12 ppm is exceeded at any approved monitoring site in the region more than 3 times in 3 years

Does this rule protect the people?

12

Rainfall measurement

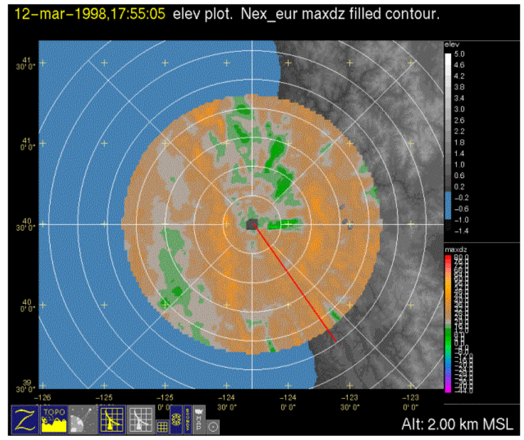
Rain gauge (1 hr)
- High wind, low rain rate (evaporation)
- Spatially localized, temporally moderate

Radar reflectivity (6 min)
- Attenuation, not ground measure
- Spatially integrated, temporally fine

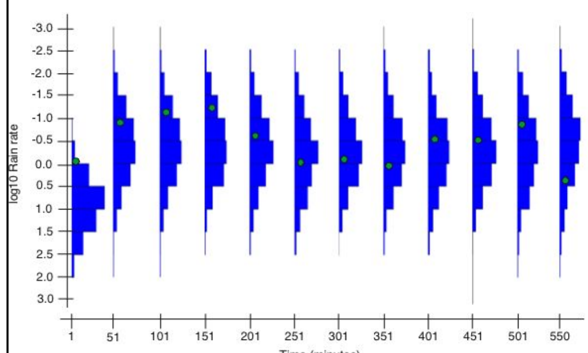
Cloud top temp. (satellite, ca 12 hrs)
- Not directly related to precipitation
- Spatially integrated, temporally sparse

Distrometer (drop sizes, 1 min)
- Expensive measurement
- Spatially localized, temporally fine

Radar image



Observed and predicted rain rate



Hematopoiesis

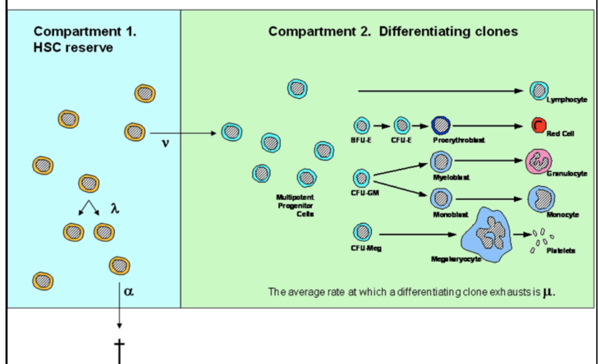
Hematopoiesis is the production of blood cells
Hematopoietic stem cells (HSC) are capable of:

- **Self-renewal**
- **Differentiation into progenitor cells**

HSCs cannot be directly observed
HSCs are defined by functional capacity

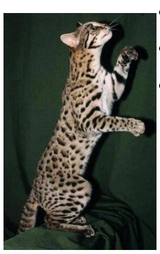
17

The model



The experiment

Safari cat (cross of Geoffroy South American wild cat and domestic cat)



- **Remove bone marrow**
- **Irradiate**
- **Replace some bone marrow**

19

Results

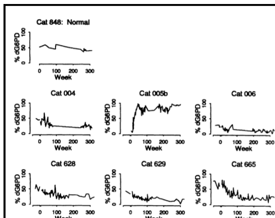
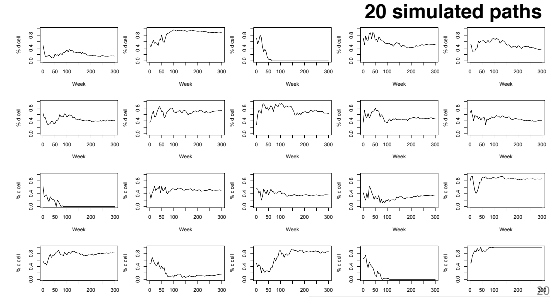


Figure 2. Observed percentages of d G6PD for one normal and six transplant cats. Week 0 represents birth in the first plot and time of transplantation in the others.

20 simulated paths



20

Probability theory

EXPERIMENT

Repeatable (controlled conditions)

Measurable well-defined outcomes

In any repetition there is always uncertainty as to which of the possible outcomes will occur

SAMPLE OUTCOME

s denotes a possible outcome of the experiment

SAMPLE SPACE

S denotes the set of sample outcomes (also called the outcome set)

EVENT

A is a subset of S; a description of the result of the experiment

21

Some experiments

Experiment: Roll a die and note which face is showing.

Sample space:

Event: Even outcome =

Experiment: A hospital has two sources of emergency power. We observe which (if any) of the emergency sources are functioning in the case of a power failure.

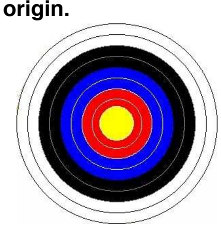
Sample space:

Event: The hospital has power during a power failure =

22

Archery

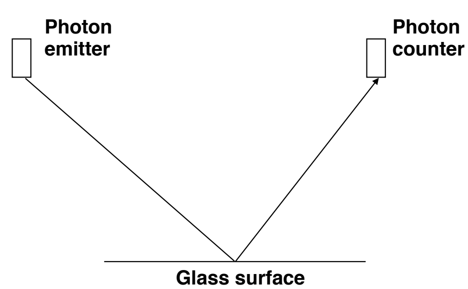
Experiment: Determine the coordinates of the arrow on the target of an archer's shot, letting the center of the target be the origin.



Sample space:

Event: Hitting the bull's eye =

23



Experiment: Emit single photons in the direction of the glass surface, and observe whether or not the photon counter advances (we call this a *hit*). Repeat this n times.

Sample space: $S = \{0,1\}^n$

$s =$ n-digit binary number

Event: k hits

24

Some set theory

The *complement* of a set A is

$$A^c = \{s \in S : s \notin A\}$$

The *union* of A and B is

$$A \cup B = \{s \in S : s \in A \text{ or } s \in B\}$$

The *intersection* of A and B is

$$A \cap B = \{s \in S : s \in A \text{ and } s \in B\} \equiv AB$$

A and B are *mutually exclusive* if

$$A \cap B = \varnothing$$

25

Examples

A dashboard warning light is supposed to flash if the car's oil pressure is too low. We are interested in the event
E = "warning light flashes when oil pressure is too low"

S =

A = Oil pressure too low =

B = Warning light flashes =

E =

26

Let S be the plane and define the events
A = { x,y l 0 < x <3, 0 < y < 3 }
B = { x,y l 2 < x < 4, 2 < y < 4 }

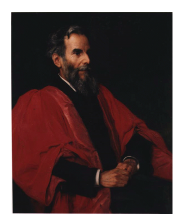
What is $A \cup B$?

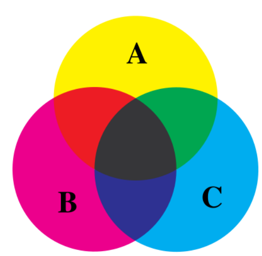
$A \cap B$?

27

Venn diagrams

John Venn (1834-1923)
British logician
1880 paper

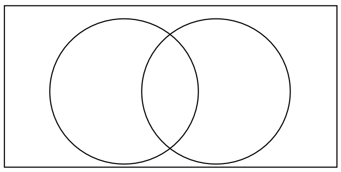




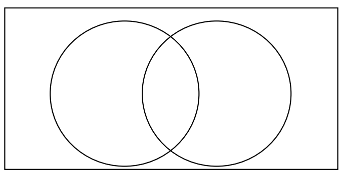
28

Examples

Precisely one of two events happen



At most one of two events happen



de Morgan's law

Augustus de Morgan
1806-1871
British mathematician

$$\left(\bigcup_{i=1}^{n} E_i\right)^c = \bigcap_{i=1}^{n} E_i^c$$

$$\left(\bigcap_{i=1}^{n} E_i\right)^c = \bigcup_{i=1}^{n} E_i^c$$

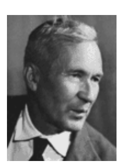
30

Probability axioms

Probability is a function from S to the reals such that
(1) P(A) ≥ 0 for all subsets A of S
(2) P(S) = 1
(3) For mutually exclusive A and B

$$P(A \cup B) = P(A) + P(B)$$

Alexei Kolmogorov
1903-1987
Foundations of the Theory of Probability
(1933)



31

Summary of last lecture

Experiment
Outcomes
Sample space
Events
Venn diagrams
Probability axioms

Office hours in B302:
Bailey F 9-11
Peter Th 8:30-10:30

32

Consequences

If $A \subseteq B$ then $P(A) \le P(B)$

$0 \le P(A) \le 1$

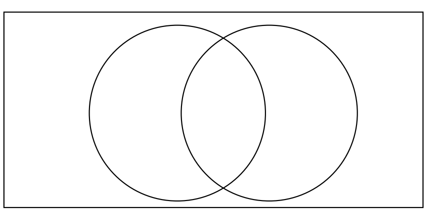
33

Complements

$P(A^c) = 1 - P(A)$

34

Unions



$P(A \cup B) = P(A) + P(B) - P(A \cap B)$

35

Example

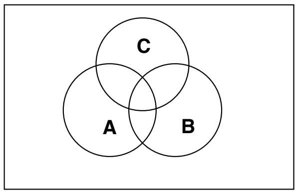
Consolidated Industries are under pressure to eliminate its discriminatory hiring practices. During the next five years they intend to hire

- **60% females**
- **30% non-white**
- **25% white males**

	N	W
F		
M		

36

Three events



$P(A \cup B \cup C) =$

37

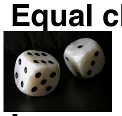
A generalization

$$P(E_1 \cup E_2 \cup \cdots E_n)$$
$$= \sum_{i=1}^{n} P(E_i)$$
$$- \sum_{i_1<i_2} P(E_{i_1} E_{i_2})$$
$$+ \cdots$$
$$+ (-1)^{r+1} \sum_{i_1<i_2<\cdots i_r} P(E_{i_1} E_{2} \cdots E_{i_r})$$
$$+ \cdots$$
$$+ (-1)^{n+1} P(E_1 E_2 \cdots E_n)$$

38

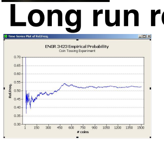
The meaning of "probability"

Equal chances

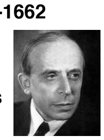


Blaise Pascal
1623-1662

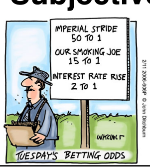
Long run relative frequency



Richard von Mises
1883-1953



Subjective probability



Bruno de Finetti
1906-1985

39

Snoqualmie Falls

Using January data from 1948-1983 at Snoqualmie Falls we find
P (rain) = 0.72 (ranging from 0.43 to 0.97 over years)

For all days that follow a rainy day we have
P (rain) = 0.83

For all days for which it rained the previous year
P (rain) = 0.73

40

Summary of Monday

Proved some probability formulas

$P(A^c) = 1 - P(A)$

$P(A \cup B) = P(A) + P(B) - P(A \cap B)$

Proved some theorems

If $A \subseteq B$ then $P(A) \leq P(B)$

$0 \leq P(A) \leq 1$

Talked about interpretations of probability

Midterm?

41

Conditional probability

Define $P(A|B) = \dfrac{P(A \cap B)}{P(B)}$

for B such that $P(B) > 0$.

Fact: $P(A|B)$ is a probability on the sample space B.

Corollary: $P(A \cap B) = P(A|B)P(B)$

42

The older child

Pick a family at random among those with two children.
Let B = {both children are girls}
P(B) =

What is the conditional probability of B, given there is at least one girl?

What is the conditional probability of B, given that the older child is a girl?

43

Total probability

Let $A_1,\ldots,A_n$ be a *partition* of S, i.e.
$A_i \cap A_j = \varnothing$, $P(A_i) > 0$ and $\bigcup_{i=1}^n A_i = S$
Then

$$P(B) = \sum_{i=1}^{n} P(B|A_i)P(A_i)$$

44

Let’s Make a Deal

Game

If the contestant picks door k, what is the probability that the car is behind door i given that door j is opened?

A con game

Three cards are put into a hat. One is white on both sides, one is red on both sides, and one is red on one and white on the other side. A card is drawn from the hat and placed on a table, without showing the underside. Suppose the top of it is red. The con man offers to bet even money on the bottom also being red. Why should you not take the bet?

Last Friday

Showed that conditional probability is probability on a new sample space
Proved the theorem of total probability
Examples
Problems

Solutions to problems

1. 3,3,3 can only come out in one way, 3,3,4 in three ways (think of the dice as different color). All the other outcomes can be matched up so there are equally many outcomes (6 when all the dice are different, 3 when two dice are the same). In total there are 27 outcomes that yield 10 and 25 that yield 9 (out of 216). (Galileo, ~1640)

2. (a) .25+.05+.05 = .35
(b) .04+.10+.05+.04+.05+.03 = .31
(c) .25/(.25+.04+.10) = .64
(d).25/(.25+.05+.05) = .71

3. (a) $P(A_k)=c/k$, and $A_1=S$ so c=1. $A_k = \{s_k\} + A_{k+1}$ so $P(A_k) = P(\{s_k\}) + P(A_{k+1})$, or $P(\{s_k\}) = 1/k - 1/(k+1)$ [= 1/(k(k+1))].
(b)

$$P(B) = \sum_{j=0}^{\infty} P(\{s_{2j+1}\}) = \sum_{j=0}^{\infty}\left[\frac{1}{2j+1} - \frac{1}{2j+2}\right]$$
$$= 1 - \tfrac{1}{2} + \tfrac{1}{3} - \tfrac{1}{4} + \cdots = \ln 2 = .69$$

4. P(A) = 2/6 P(B) = 6/36 P(AB) =2/36 = 2/6 x 6/36 so independent

5.

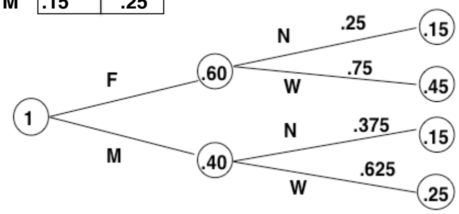
$$P(O \mid N) = \frac{P(O \,\&\, N)}{P(O)} = \frac{P(N \mid O)P(O)}{P(N \mid O)P(O) + P(N \mid O^c)P(O^c)}$$
$$= \frac{.8 \times .65}{.8 \times .65 + .2 \times .35} = .88$$

6. (a) a x .5 + b x .5
(b) P(claims two years in a row) = a^2 x .5 + b^2 x .5 so the probability sought is $(a^2+b^2)/(a+b)$
(c) a / (a+b)

Probability trees

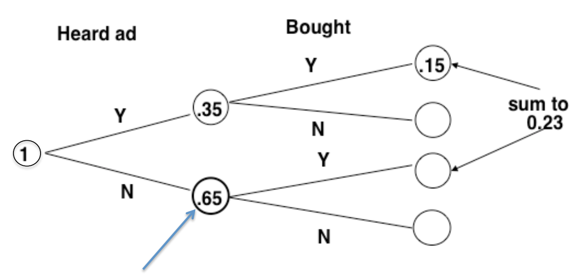
A *probability tree* is a picture indicating probabilities and conditional probabilities for combinations of two or more events. Consolidate Industries:

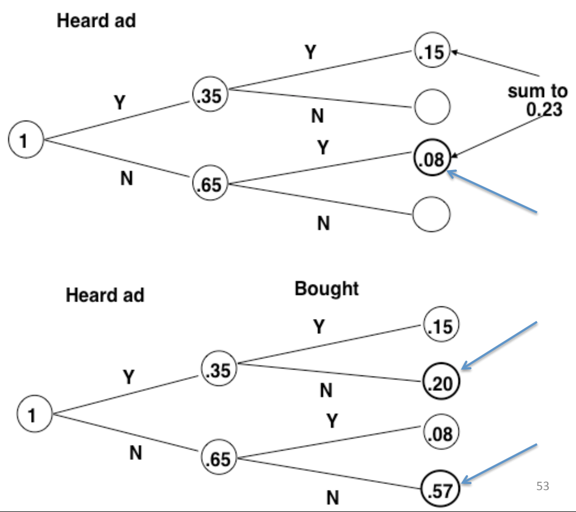
	N	W
F	.15	.45
M	.15	.25



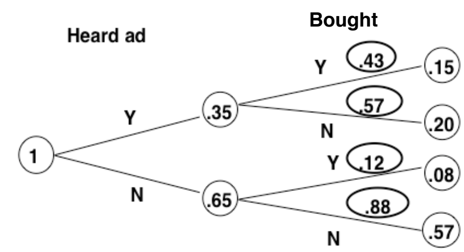
Constructing the tree

We are told the following three probabilities:
P(heard the advertisement) = 0.35
P(bought the product) = 0.23
P(heard the ad and bought the product) = 0.15





53



P(bought I heard) = 0.15 / 0.35 = 0.43.

Conversely, P(heard I bought) = 0.15 / (0.15+0.08)
= 0.65

54

Rules for probability trees

Probabilities add to one at each level of the tree

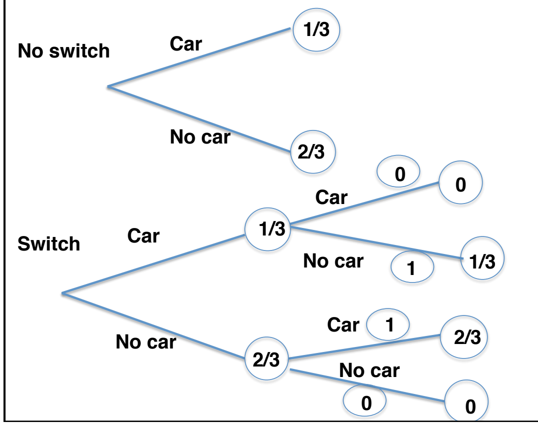
Conditional probabilities add to one for each group of branches radiating from a single point

The probability at a branch point times the conditional probability along the branch is the probability at the end of the branch

The probability at a branch point is the sum of all probabilities at the end of branches to the right of the point

55

Let's make a deal, revisited



56

Bayes' theorem

Let $A_1,\ldots,A_n$ be a partition of S.
Assume P(B) > 0. Then

$$P(A_i|B) = \frac{P(B|A_i)P(A_i)}{\sum_{j=1}^{n} P(B|A_j)P(A_j)}$$

57

Screening for cervical cancer

A = {positive Pap smear}
B = {cervical cancer}

P(B)=0.000081 (2002-2005)
P(A|B) = 0.72
P(A|B^c) = 0.06

$$P(B|A) = \frac{0.000058}{0.0601} = 0.0010$$

What is the error rate for the Pap smear?

58

Monday's lecture

Solutions to Friday's problem set
Probability trees
Bayes' formula

59

Asking sensitive questions

How can you get accurate answers to questions about drug use, cheating etc.

A: Is your student number even?
B: Have you ever cheated at an exam at the University of Washington?
Flip a coin. Heads answer A, tails answer B. DO NOT SAY which question you are answering!
P(even student number) = 24/46
Result: Yes 18 No 20 Invalid 1

60

US Census data

Race	Male	Female	All
White	1,487	1,412	2,899
Nonwhite	366	348	714
Total	1,853	1,760	3,613

Data on newborns from 1980 US census in thousands of individuals. Consider a randomly chosen baby.

P(female) =

P(female I nonwhite) =

61

Independence

A and B are independent if P(A|B) = P(A).
Equivalently $P(A \cap B) = P(A)P(B)$

62

Question

Suppose that two events are mutually exclusive. Are they independent?

63

Genetics

Suppose a certain trait is determined by a gene with two alleles: a *dominant* W and a *recessive* w. Let population frequencies of WW be p, of Ww be q, and of ww be r. If we assume *random mating* (parents selected randomly and a randomly and independently chosen allele transferred from each parent), what are the corresponding frequencies among the offspring?

Let A_w = receive w from father

B_w = receive w from mother

Then $P(\text{offspring is ww}) = P(A_w \cap B_w)$

64

Genetics, cont.

By total probability

$P(A_w) = P(A_w \mid WW)\ P(WW)$

$+ P(A_w \mid Ww)\ P(Ww)$

$+ P(A_w \mid ww)\ P(ww)$

$= 0 \times p + 1/2 \times q + 1 \times r$

$= r + q/2 = P(B_w)$

65

Genetics, cont.

By the independence assumption the probability of the offspring receiving the recessive trait, i.e., genotype ww, is

$$P(A_w \cap B_w) = P(A_w)P(B_w) = (r + q/2)^2$$

How about WW?

One of each allele?

These probabilities are called the *Hardy-Weinberg* proportions.

66

More than two events

We can't define independence between three events by pairwise independence, or by saying

$$P(A \cap B \cap C) = P(A)P(B)P(C)$$

We need this:

$A_1,\ldots,A_n$ are independent if for all distinct $i_1,\ldots,i_k$ with $1 \le i_j \le n$

$$P(A_{i_1} \cap \cdots \cap A_{i_k}) = P(A_{i_1}) \cdots P(A_{i_k})$$

67

Breathing Ceasar's last breath

The number of molecules in the earth's atmosphere is about 10^{44}. The molecules present today are pretty much the same as those present thousands of years ago. A typical breath consists of about 10^{22} molecules. What is the chance that your next breath will contain at least one molecule from Caesar's last breath?

68

Friday's lecture

Randomized response survey
Independence
Examples
Problems

Based on Bailey's survey I have moved my office hours to Thursday 1:30-3:20 pm in B302. Bailey's stay the same (Friday 9-11 in B302).

Problem solutions

1. The easiest way to do this is with Venn diagrams.
   Answers: T F F F T T T
2. (a) No. Let A=C with P(A) = .5
   (b) No. Consider tossing two fair coins. A="heads on first", C="heads on second", B="same outcome on both"

$$P(B) = 1/4 + 1/4 = 1/2$$

$$P(A \cap B) = P(A \cap C) = 1/4 = P(A)P(B)$$

so A and B are independent, as are A and C.

$$P(A \cup C) = P(A) + P(C) - P(A \cap C) = 3/4$$

$$P(B \cap (A \cup C)) = P(A \cap C) = 1/4 \neq 1/2 \times 3/4$$

(c) No. Same setup as (b).

$$P(B \cap A \cap C) = P(A \cap C) = 1/4 \neq 1/2 \times 1/4$$

3.(a) If B attracts A then

$$P(A \cap B) > P(A)P(B)$$

so A attracts B. Now write

$$P(A|B^c)P(B^c) = P(A \setminus B) = P(A) - P(A \cap B)$$

$$= P(A) - P(A|B)P(B) < P(A) - P(A)P(B)$$

$$= P(A)P(B^c)$$

(b) Since $P(F_j \mid B_i) = 1$ for $i \neq j$ Bayes' theorem yields

$$P(B_j|F_j) = \frac{P(F_j|B_j)P(B_j)}{\sum_{i=1}^{n} P(F_j|B_i)P(B_i)} = \frac{f_j b_j}{1 - b_j + f_j b_j}$$

$$P(B_j) - P(B_j|F_j) = \frac{b_j(1-b_j)(1-f_j)}{1 - f_j + b_j f_j} > 0$$

and for $i \neq j$

$$P(B_i|F_j) - P(B_i) = \frac{b_i}{1 - b_j - b_j f_j} - b_i = \frac{b_i b_j (1-f_j)}{1 - b_j - b_j f_j} > 0$$

4. Let A be the event that A tells the truth, etc., and S={A makes given statement}. First if A is true then

$$D \cap C \cap B \cap A \subseteq S$$

and

$$P(D \cap C \cap B \cap A) = 1/81$$

Also if D then two of A,B,C must lie, and each of these three possibilities has probability 4/81. When D lies, they can all lie, 16/81, or one of A,B,C can lie, each 4/81. These are all the options, so

$$P(D|S) = \frac{P(D \cap S)}{P(D \cap S) + P(D^c \cap S)} = \frac{\frac{13}{81}}{\frac{13}{81} + \frac{28}{81}} = \frac{13}{41}$$

73

5. A={0 recieved}, B={0 sent}

Sent	Relay 1	Relay 2	Relay 3	Probability
0	0	1	0	3/64
0	1	1	0	3/64
0	1	0	0	3/64
0	0	0	0	27/64

$$P(A|B) = \frac{(36/64)p}{(36/64)p + (28/64)(1-p)}$$

where p is the frequency of 0.

6. A={6 on first} B= {6 on second}
F={fair die picked} L={loaded die}

$$P(A \cap B) = P(A \cap B|F)P(F) + P(A \cap B|L)P(L)$$

$$= \frac{1}{36} \times \frac{3}{9} + \frac{1}{4} \times \frac{6}{9} = \frac{19}{108}$$

74

$$P(A) = P(B) = \frac{1}{6} \times \frac{3}{9} + \frac{1}{2} \times \frac{6}{9} = \frac{7}{18}$$

$$P(A)P(B) = \frac{49}{324} \neq \frac{57}{324} = P(A \cap B)$$

so the two events are not independent.

75

People vs. Collins

1964 LA purse snatching by a young blond female with ponytail, escaping with a black male with moustache and beard in a yellow car. A suspect with these characteristics was found. The state argued that it would be very unlikely that there would be more than one such couple.

Characteristic	Probability
Female with ponytail	1/10
Yellow automobile	1/10
Man with moustache	1/4
Blond female	1/3
Black male with beard	1/10
Interracial couple	1/1000

The product is 1/12 million.

Bertillon system of identification

Predates fingerprints

11 anatomical variables, such as height, head width, ear length, which do not change much for adults

Each classified as small, medium, or large: e.g.,

(s, m, m, s, l, s, m, m, s, s, m)

How likely would it be that in Seattle there are two individuals with the same Bertillon configuration?

77