ORF 245 Fundamentals of Statistics
Chapter 1
Probability

Robert Vanderbei

Sept 2014

Slides last edited on September 19, 2014

Course Info

Prereqs: Three semesters of Calculus

Textbook: *Mathematical Statistics and Data Analysis* by John A. Rice

Grading:

Homework:	45%
Midterm 1:	15%
Midterm 2:	15%
Final:	20%
Participation:	5%

Homework:

- Will be due every week at 5pm on Friday.
- Turn homework in via ORF245 drop box in Sherrerd Hall.
- The lowest homework grade will be dropped.

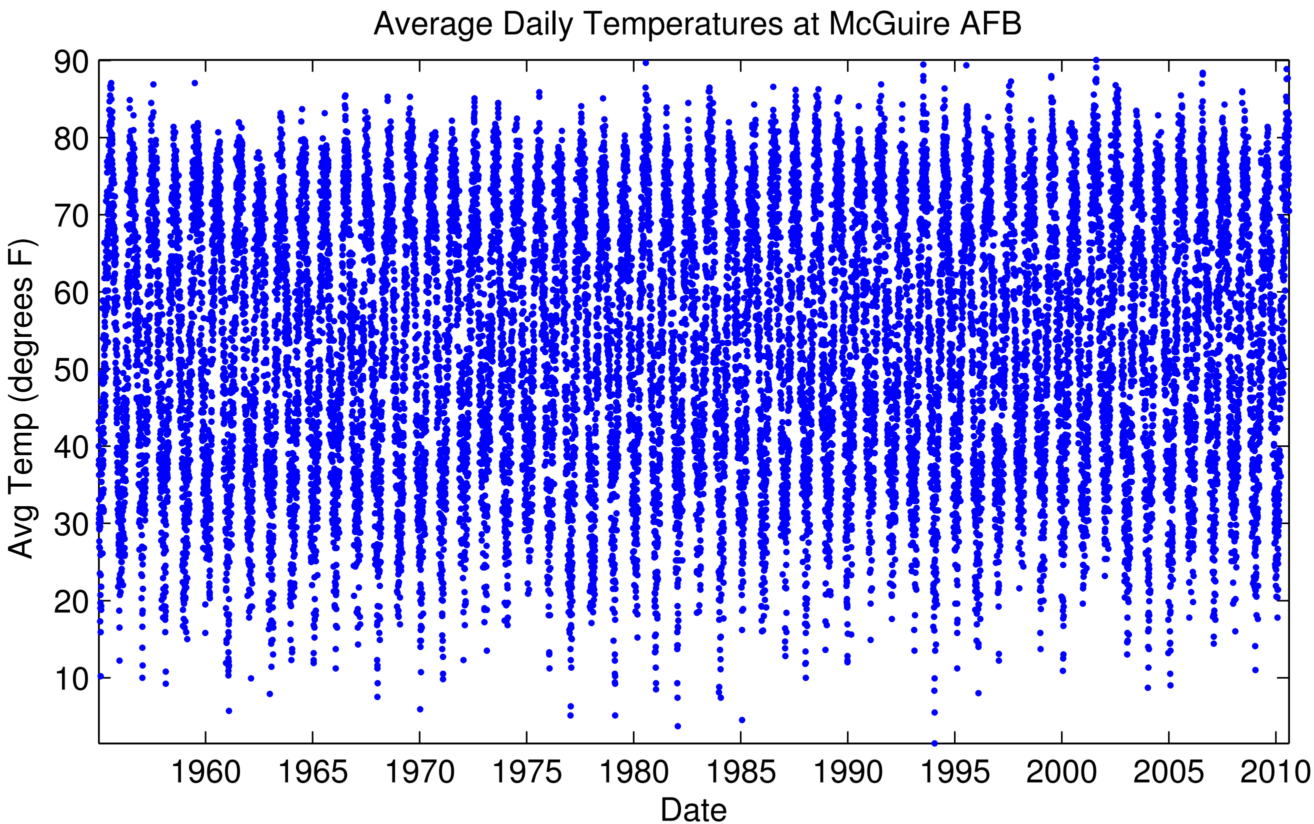
Midterms: Midterms will be in-class on Wednesday of the 5th and 10th weeks.

Lectures: Reading material will be posted in advance of each lecture. You are expected to read the reading material before lecture.

Slides: The slides will be posted online. But, they are not a replacement for the lecture. They are just my notes to remind me what to say. You must go to lecture to hear what I have to say.

Introduction

Statistics is about extracting meaningful conclusions from noisy data. Here's an example.



Is there a warming trend? If yes, what is the rate of warming?
Eventually, we will answer this question.

Matlab

In this class we will do some statistical computing.

We will also want to be able to plot data.

For these tasks, we will use a computer programming language called *Matlab*.

Here's the code that was used to make the plot on the previous page.

```
load -ascii 'mat_output'
t=mat_output(:,1);
avg=mat_output(:,2);
plot(t,avg,'.b');
xlabel('Date');
ylabel('Avg Temp (degrees F)');
title('Average Daily Temperatures at McGuire AFB');
```

Before delving deeply into statistics, we need to have a good grasp of what we mean by "noisy" data.

So, we lay the groundwork with some probability...

Sample Spaces and Events

When considering experiments *to be performed*, we need to define the *set of possible outcomes*.

The set is called the *sample space* and is usually denoted by Ω.

A sample space can be either finite or infinite.

Examples:

- Toss a coin: $\Omega = \{H, T\}$.
- Roll a die: $\Omega = \{1, 2, 3, 4, 5, 6\}$.
- Roll a pair of dice: $\Omega = \{(i, j) \mid 1 \leq i \leq 6, 1 \leq j \leq 6\}$.
- How long from now until the next major earthquake in California: $\Omega = \{t \mid t \geq 0\}$.

Subsets of Ω are called *events*. They are usually denoted by capital letters like A or B.

Example:

- Roll a pair of dice and consider the event that the sum is 4:

$$A = \{(1, 3), (2, 2), (3, 1)\} \subset \Omega.$$

And, Or, Not, etc.

Intersection:	Event A *and* B both occur is written $A \cap B$.
Union:	Event A *or* B occurs is written $A \cup B$.
Complement:	Event A does *not* occur is written A^c.
Empty Set:	The set containing *no elements* is denoted $\emptyset$.
Disjoint:	Sets A and B are *disjoint* if $A \cap B = \emptyset$.

Probability

Probabilities are numbers assigned to events. They must satisfy the following properties:

- $P(\Omega) = 1$.
- $P(A) \geq 0$ for all $A \subset \Omega$.
- If A_1 and A_2 are disjoint, then $P(A_1 \cup A_2) = P(A_1) + P(A_2)$.

It follows that

- $P(A^c) = 1 - P(A)$.
- $P(\emptyset) = 0$.
- $P(A) \leq P(B)$ whenever $A \subset B$.

Finite Sample Spaces

Examples:

1. Flip two pennies and let $\Omega = \{(h,h),(h,t),(t,h),(t,t)\}$. Then,

$$\begin{aligned} P(\text{one head and one tail}) &= P(\{(h,t)\} \cup \{(t,h)\}) = P(\{(h,t)\}) + P(\{(t,h)\}) \\ &= \frac{1}{2}\cdot\frac{1}{2} + \frac{1}{2}\cdot\frac{1}{2} = \frac{1}{2} \end{aligned}$$

2. Flip two pennies and let $\Omega = \{$"two heads", "one head", "no heads".$\}$. Then,

$$P(\text{one head and one tail}) = \text{hmmm... hard to say}$$

3. Roll a pair of dice and let $\Omega = \{(i,j) \mid 1 \le i \le 6, 1 \le j \le 6\}$. Then,

$$P(\text{sum is } 4) = P(1,3) + P(2,2) + P(3,1) = 3/36$$

Conditional Probability

Definition. The conditional probability of A given that B is known to have occurred is

$$P(A|B) = \frac{P(A \cap B)}{P(B)}.$$

Law of Total Probability. Let $B_1, B_2, \ldots, B_n$ be a disjoint collection of sets each having positive probability whose union is all of Ω. Then,

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

Bayes' Rule. If, in addition to the assumptions above, $P(A) > 0$, then

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

If the disjoint collection consists of just two sets, B and B^c, then the formula can be written more simply as

$$P(B|A) = \frac{P(A|B)P(B)}{P(A|B)P(B) + P(A|B^c)P(B^c)}.$$

Example of Bayes' Rule

Consider women of a certain given age and overall health status.

Facts:

- Probability that a woman has breast cancer $= P(B) = 1\%$.
- Probability that a mammogram will give a positive result (indicating cancer is present) (event A) for women who are known to have cancer $= P(A|B) = 80\%$.
- A women who is known not to have cancer will test positive 10% of the time. That is, $P(A|B^c) = 10\%$.

Question: If a woman tests positive for breast cancer, what is the probability that she actually has breast cancer? That is, what is $P(B|A)$?

Cancer doctors were asked this question. Most estimated the answer to be 75%.

Example of Bayes' Rule

Consider women of a certain given age and overall health status.

Facts:

- Probability that a woman has breast cancer $= P(B) = 1\%$.
- Probability that a mammogram will give a positive result (indicating cancer is present) (event A) for women who are known to have cancer $= P(A|B) = 80\%$.
- A women who is known not to have cancer will test positive 10% of the time. That is, $P(A|B^c) = 10\%$.

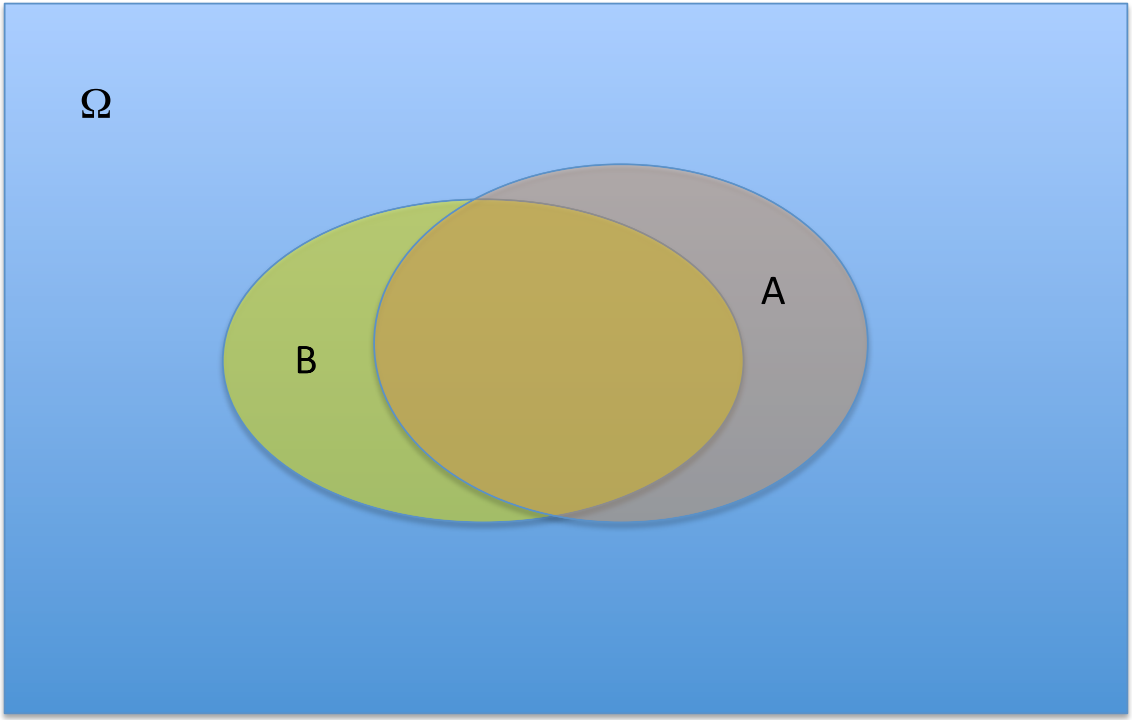
Question: If a woman tests positive for breast cancer, what is the probability that she actually has breast cancer? That is, what is $P(B|A)$?

Cancer doctors were asked this question. Most estimated the answer to be 75%.

Let's compute:

$$P(B|A) = \frac{(0.8)(0.01)}{(0.8)(0.01) + (0.1)(0.99)} = \frac{8}{8 + 99} \approx 7.5\%.$$

Venn Diagram



Independence

Events A and B are independent means

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

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

Independence

Events A and B are independent means

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

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

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

Ch. 1, Number 13

In a game of poker, what is the probability that a five-card hand will contain

(a) a straight (five cards in a numerical sequence not all from same suit),

(b) four of a kind (four cards of one value), and

(c) a full house (three cards of one value and two cards of another)?

(a)

$$\frac{10 \cdot (4^5 - 4)}{(52 \cdot 51 \cdot 50 \cdot 49 \cdot 48)/(5 \cdot 4 \cdot 3 \cdot 2 \cdot 1)} = 0.00392$$

(b)

$$\frac{13 \cdot 48}{\binom{52}{5}} = 0.000240$$

(c)

$$\frac{13 \cdot 12 \cdot \binom{4}{3}\binom{4}{2}}{\binom{52}{5}} = 0.001441$$

Example E: Birthday Problem

In a classroom of n students, what's the probability p_n that two (or more) students share the same birthday?

It's easier to compute the probability that no two students share a birthday.
Let's look at the students one at a time.
The first student can have any birthday he/she likes.
The second student cannont share the first student's birthday: 364 choices.
The third student cannot share either of the first two birthdays: 363 choices.
... Etc. ...
The n-th student cannot share any of the previous $n-1$ birthdays: $365-n+1$ choices.
Therefore, probability of no shared birthdays is

$$\frac{365 \cdot 364 \cdots (365-n+1)}{365^n}$$

and the probability of a shared birthday is

$$p_n = 1 - \frac{365 \cdot 364 \cdots (365-n+1)}{365^n}.$$

For $n = 23$, the answer is

$$p_{23} = 0.507.$$

Example: Monty Hall Problem

Suppose you're on a game show, and you're given the choice of three doors: Behind one door is a car; behind the others, goats. You pick a door, say No. 1, and the host, who knows what's behind the doors, opens another door, say No. 3, which has a goat. He then says to you, "Do you want to pick door No. 2?" Is it to your advantage to switch your choice?

Ch. 1, Number 56

A couple has two children.

(a) What is the probability that both are girls given that the oldest is a girl?

(b) What is the probability that both are girls given that one of them is a girl?

Let $\Omega = \{(f,f),(f,m),(m,f),(m,m)\}$ (gender of older followed by gender of younger).

Let $A =$ "both are girls" $= \{(f,f)\}$.

Let $B =$ "oldest is a girl" $= \{(f,f),(f,m)\}$.

Let $C =$ "at least one is a girl" $= \{(f,f),(f,m),(m,f)\}$.

(a) $P(A|B) = 1/2$.

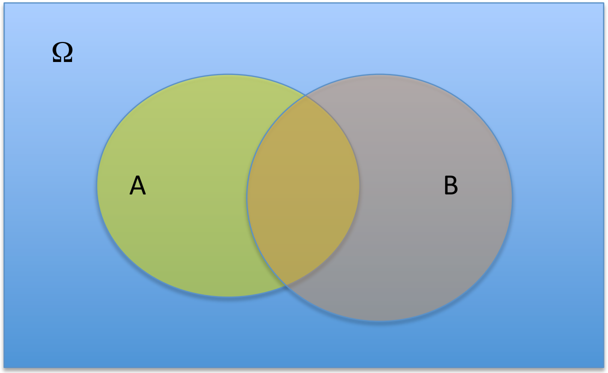
(b) $P(A|C) = 1/3$.

Show that if A and B are independent, then A and B^c are independent and so are A^c and B^c.

Given: $P(A \cap B) = P(A)P(B)$.

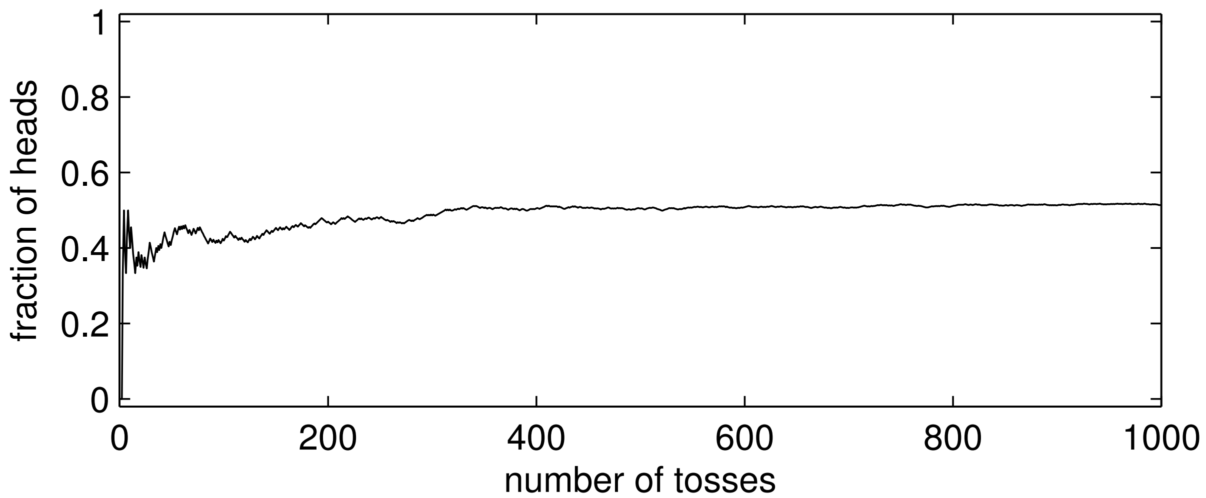
Compute:

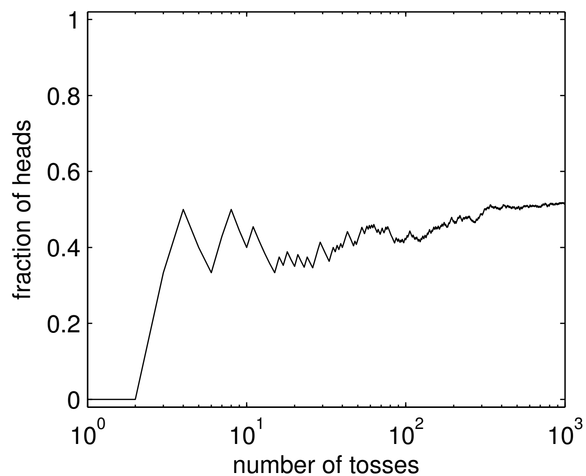
$$\begin{aligned}
P(A \cap B^c) &= P(A) - P(A \cap B) \\
&= P(A) - P(A)P(B) \\
&= P(A)\left(1 - P(B)\right) \\
&= P(A)P(B^c)
\end{aligned}$$

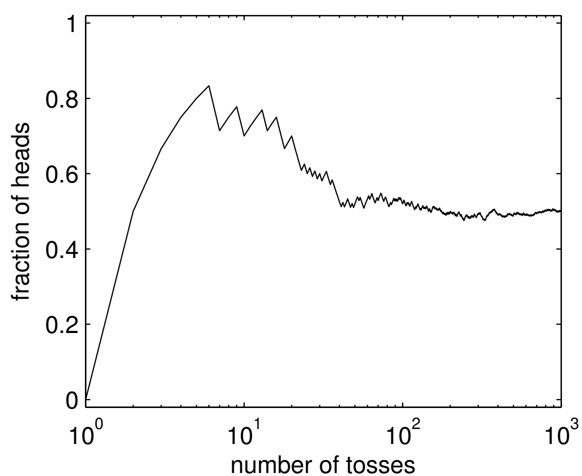


$$\begin{aligned}
P(A^c \cap B^c) &= P((A \cup B)^c) \\
&= 1 - P(A \cup B) \\
&= 1 - P(A) - P(B) + P(A \cap B) \\
&= 1 - P(A) - P(B) + P(A)P(B) \\
&= \left(1 - P(A)\right)\left(1 - P(B)\right) \\
&= P(A^c)P(B^c)
\end{aligned}$$

Frequentist Approach – Coin Tossing







Matlab Code

```
ht = randi([0,1],[1,1000]);
St = cumsum(ht);
Xt = St./(1:1000);

figure(1);
plot(1:1000,Xt,'k-');
ylim([-0.02 1.02]);
xlabel('number of tosses');
ylabel('fraction of heads');
```