AESB2440: Geostatistics & Remote Sensing

Lecture 2: Probability

Wednesday, April 22, 2015

Roderik Lindenbergh

Dept. of Geoscience and Remote Sensing

TU Delft

Delft University of Technology

Introduction Probability

A. Probability and events

- Definition(s)
- Probability/Sample space
- Set theory

B. Conditional probability

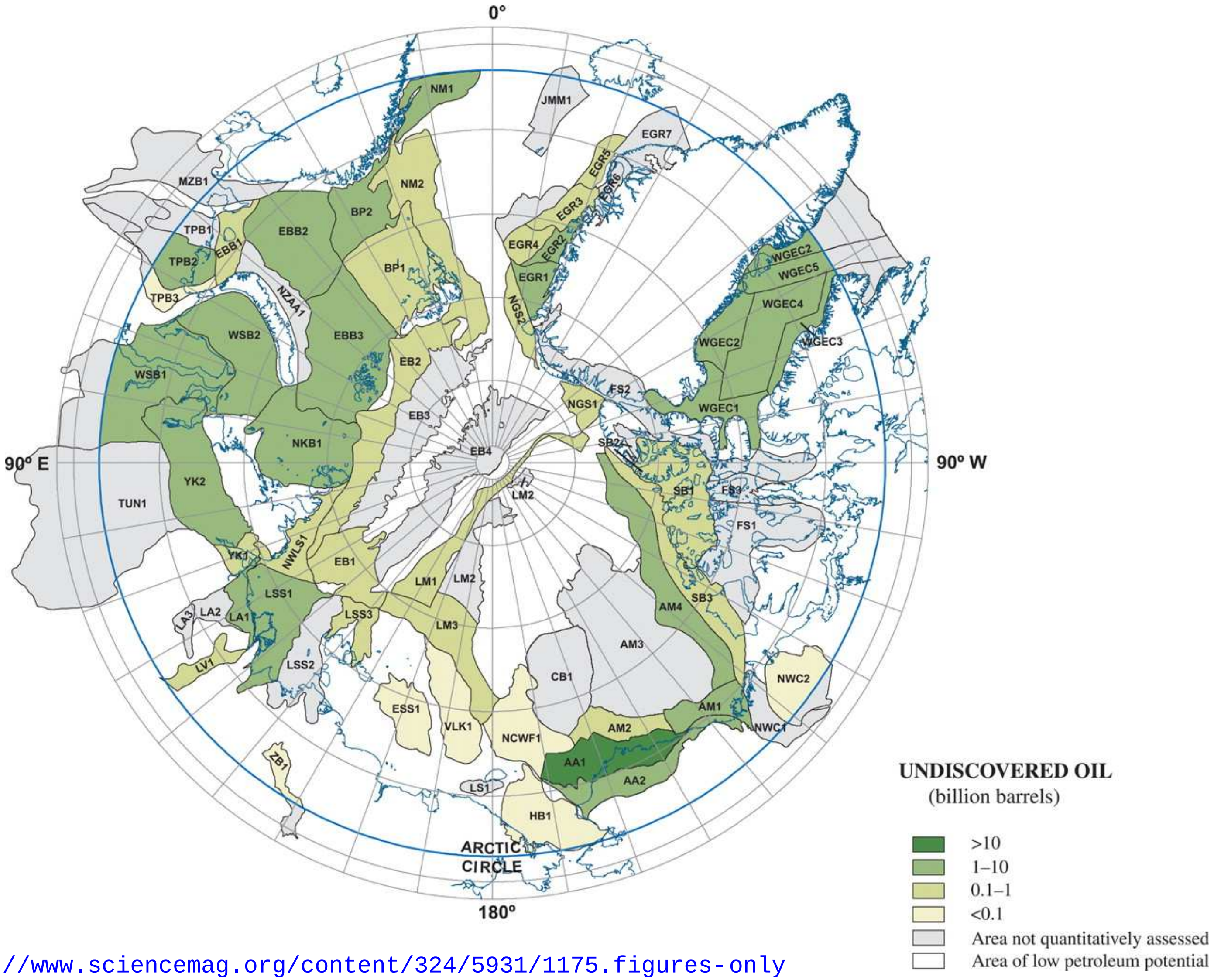
- Conditional probability
- Total probability rule
- Bayes rule
- Independency

C. Random variables & Distributions

- Random variable
- Discrete and continous RVs
- Probability mass function
- Cumulative distribution function

Why probability?

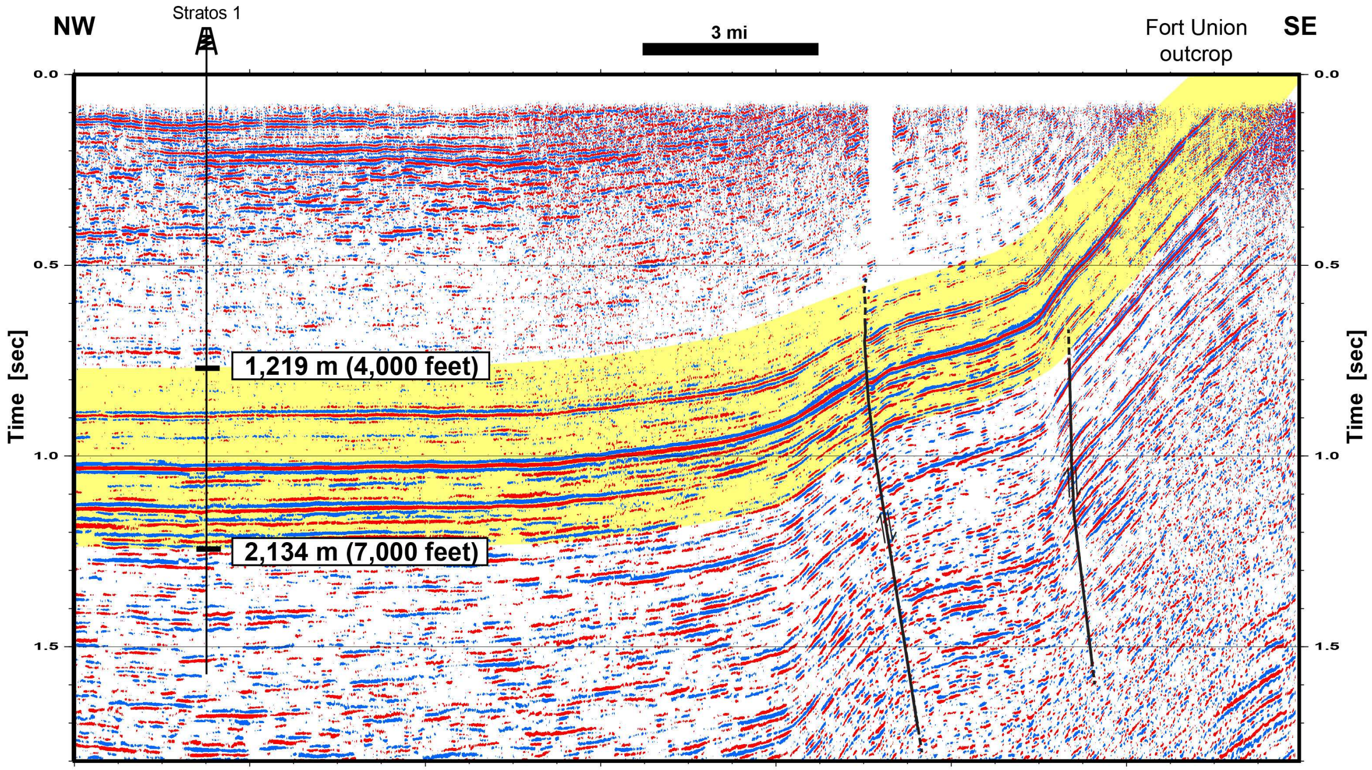
Where to look for oil?



Source http://www.sciencemag.org/content/324/5931/1175.figures-only

Oil probabilities

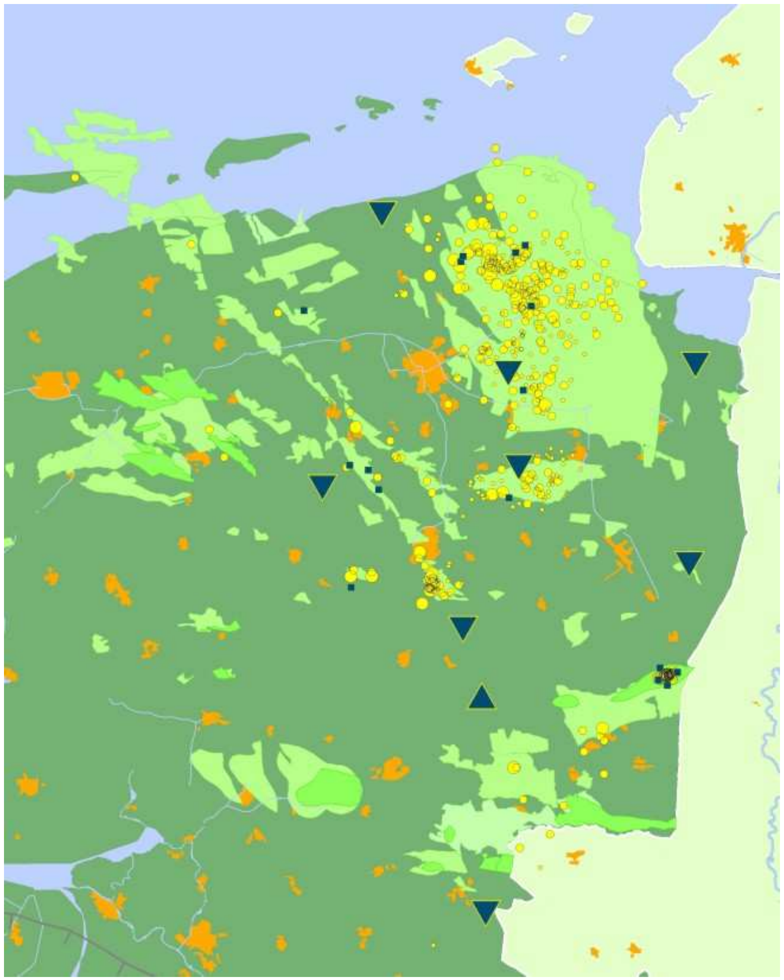
How certain should you be before you build an oil platform?

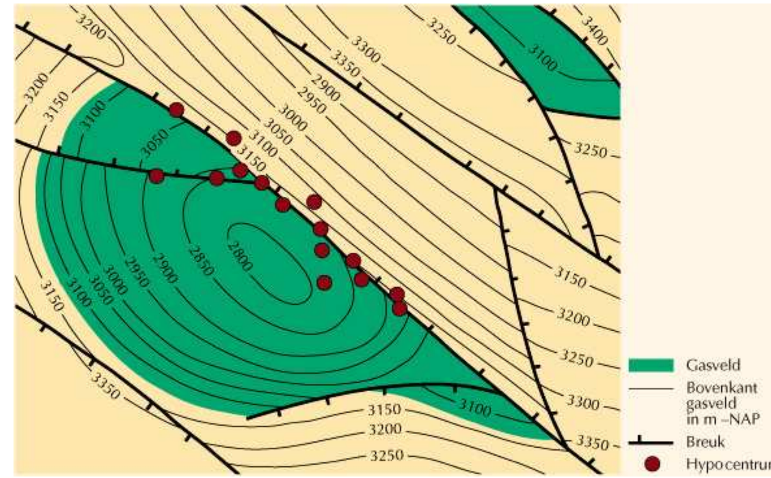


Starts with interpretation of uncertain seismic data

Source http://www.wsgs.wyo.gov/research/energy/oil-gas/hydro-carbons.aspx

Groningen Earthquakes





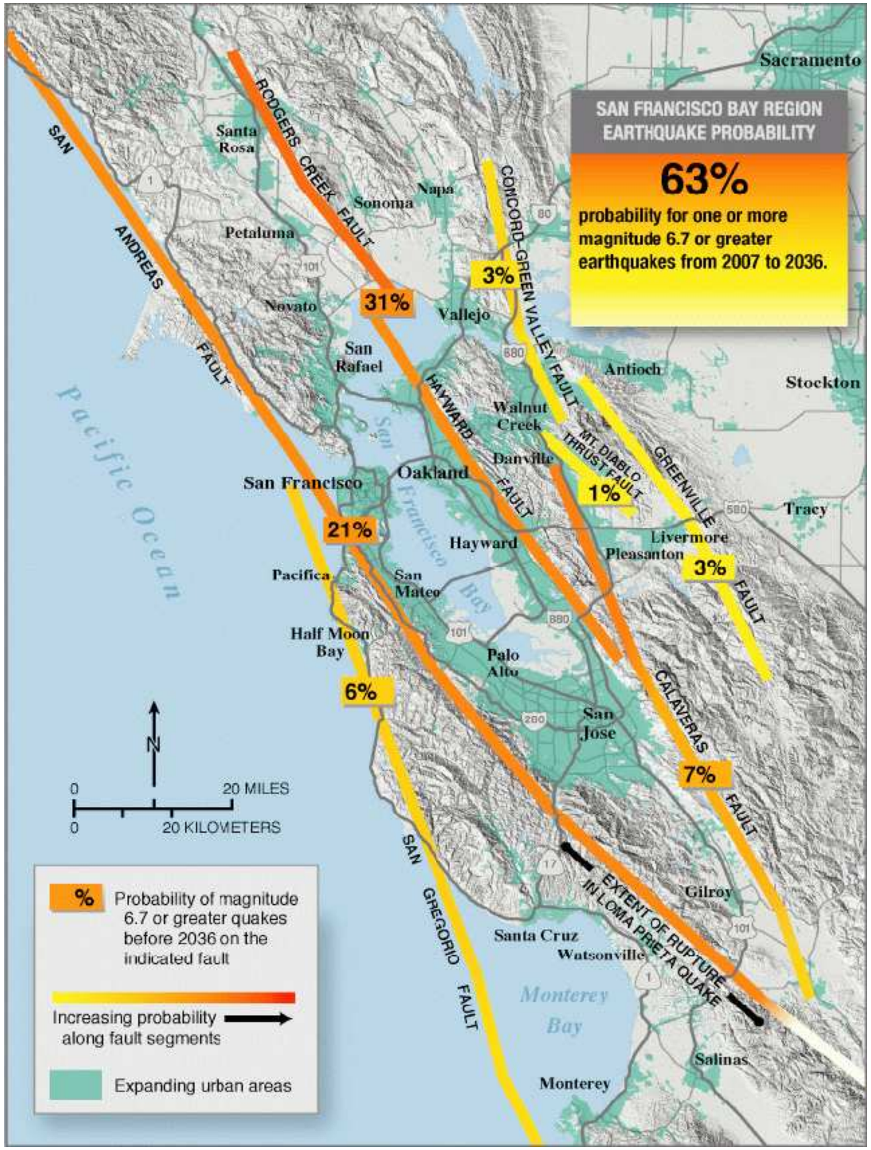
- Yellow circles: earthquakes;
- Lightgreen: gasfields;
- Blue triangles: drillhole seismometers;
- Squares: accelerometer

Source http://www.knmi.nl/

Bay area, Earthquake probability

Question.
Why are the local probabilities lower then the overall probability of 63 %?

Question.
Why are the sum of the local probabilities larger then the overall probability of 63 %?



Source http://seismo.berkeley.edu/blog/seismoblog.php/2008/10/10/earthquake-probabilities-in-the-bay-area

Earthquake questions

≈Final question: what is the probablity that my house is damaged by an earthquake within, say, the next 50 years?

Possible subquestions

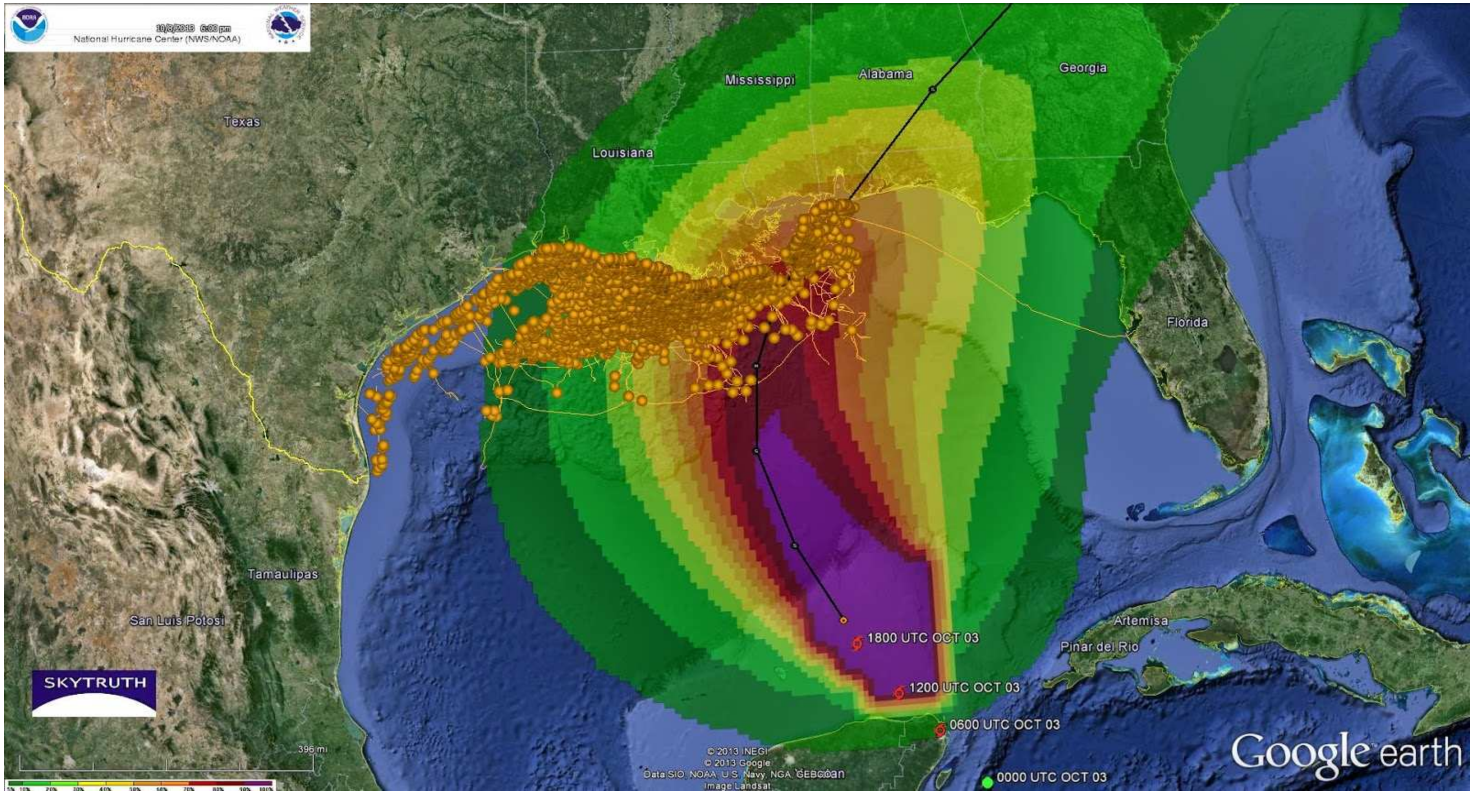
- . . .

See also: http://www.geo.mtu.edu/UPSeis/locating.html:
How Do I Locate That Earthquake's Epicenter?

Dept. of Geoscience and Remote Sensing

TU Delft

Storm development



Forecast showing probability of tropical storm-force winds occurring over the next 120 hours. Offshore oil and gas platforms are orange dots; seafloor pipelines are thin orange lines. Forecast data from NOAA/NWS/NHC.

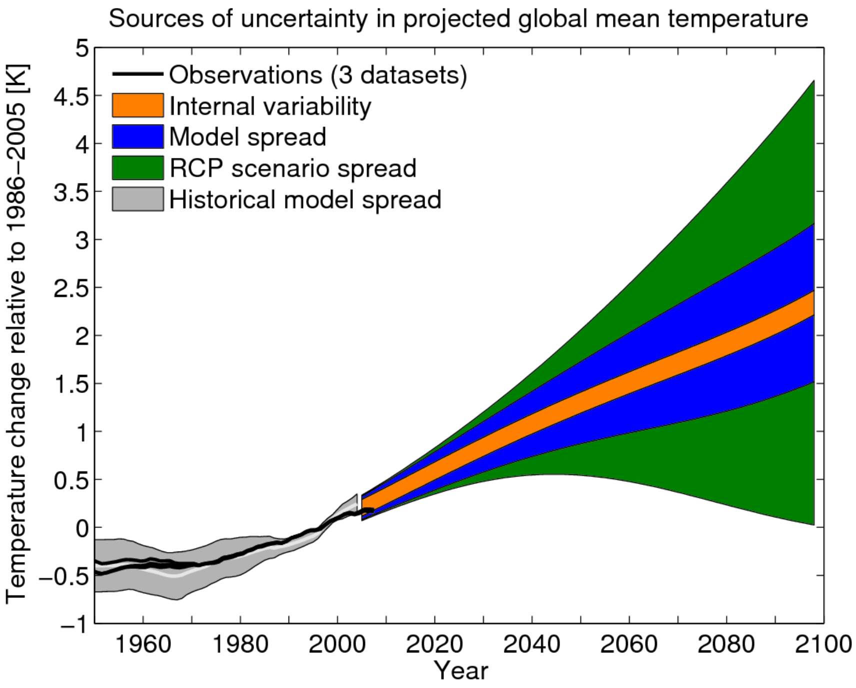
Source http://blog.skytruth.org/2013/10/tropical-storm-karen-building-in-gulf.html

Commonalities in these examples

1. Starting point:
uncertain observations

2. Modeling steps:
   - Interpolation
   - Geometry extraction
   - Complicated weather model
   - . . .

3. Quality analysis
   - Interior: how does uncertainty in the observations propagate into uncertainty of the results?
   - Ground truth: compare to (locally) available observations of better quality



Source http://www.climate-lab-book.ac.uk/2013/sources-of-uncertainty/

Dealing with measurements: requirements

Measurements are not exact: each measurement is uncertain

Requirement 1.
We want to be able to express the uncertainty of measurements

$\Rightarrow$ Conclusions based on measurements are uncertain as well.

Requirement 2.
We want to be able to propagate the uncertainty in the measurements into the conclusions

Requirement 3.
We want to be able to express the uncertainty of conclusions

Literature

1. Slides

2. Background information:

- A Modern Introduction to Probability and Statistics, Understanding Why and How
  F.M. Dekking, C. Kraaikamp, H.P. Lopuhaä, L.E. Meester,
  Springer, 2005

- Primer on Mathematical Geodesy
  C.C.J.M. Tiberius,
  Lecture Notes Surveying and Mapping, CTB3310, Dept. of Geoscience & Remote Sensing
  (Available on Blackboard)

- WWW (Wikipedia, www.mathworld.com, Google, ...)

A. Probability

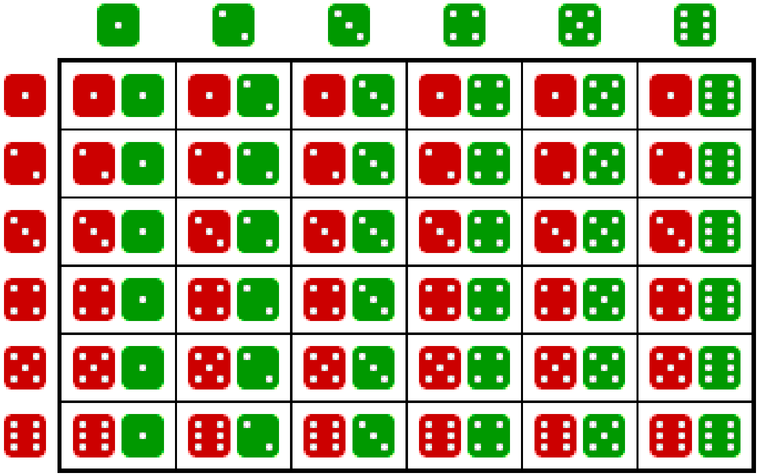
Sample or Probability space

Consider an experiment (like a measurement)

Sample space. Set of possible outcomes of an experiment. Here denoted Ω.

Experiment examples

- Tossing a coin
- Throwing a dice
- Throwing a dice twice
- Measuring somebody's length
- Order in which GRS students enter the classroom



Question. What are the corresponding sample spaces?

Source http://dsearls.org/courses/M120Concepts/ClassNotes/Probability/420$_$principles.htm

Events

SAMPLE SPACE FOR
A PAIR OF DICE

	1	2	3	4	5	6
1	2	3	4	5	6	7
2	3	4	5	6	7	8
3	4	5	6	7	8	9
4	5	6	7	8	9	10
5	6	7	8	9	10	11
6	7	8	9	10	11	12

Event.
Subset of the sample space.

Question: Is 5×5 an event for the double dice experiment?

Let Ω be the sample space, or set of all possible outcomes of an experiment

Ω Sure event (occurs always)

$\emptyset$ Impossible event (occurs never)

Source `http://kids.britannica.com/comptons/art-53854/Sample-space-for-a-pair-of-dice`

All subsets

Theorem [Power set]. If Ω contains n outcomes, there are 2^n events.

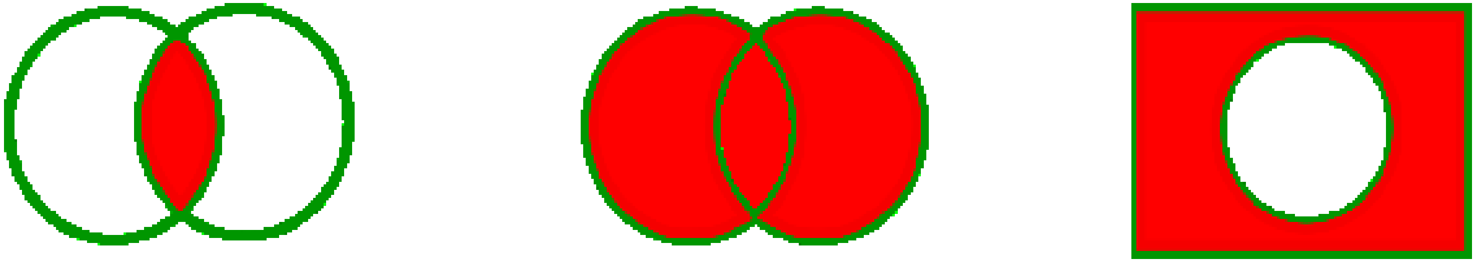
[Proof.] Binary numbers: write for each outcome a 1 if it occurs and 0 else. There are 2^n different binary ‘words‘ on n letters, where each letter corresponds to an outcome

Example. Toss a coin twice. There are $n = 4$ outcomes:

$$\{HH, HT, TH, TT\}$$

- What is the number of different events?
- What is the event $A = \{$head at first toss$\}$
- List all events

Combining events



$A \cap B$ $\qquad$ $A \cup B$ $\qquad$ A^c

$A \subset B$	subset
$A = B$	equality
$A \cap B$	intersection
$A \cup B$	union
A^c	complement

Disjoint. $A \cap B = \emptyset$, that is, A and B have no elements in common

Example. Draw a Venn diagram of $A \cap B^C$

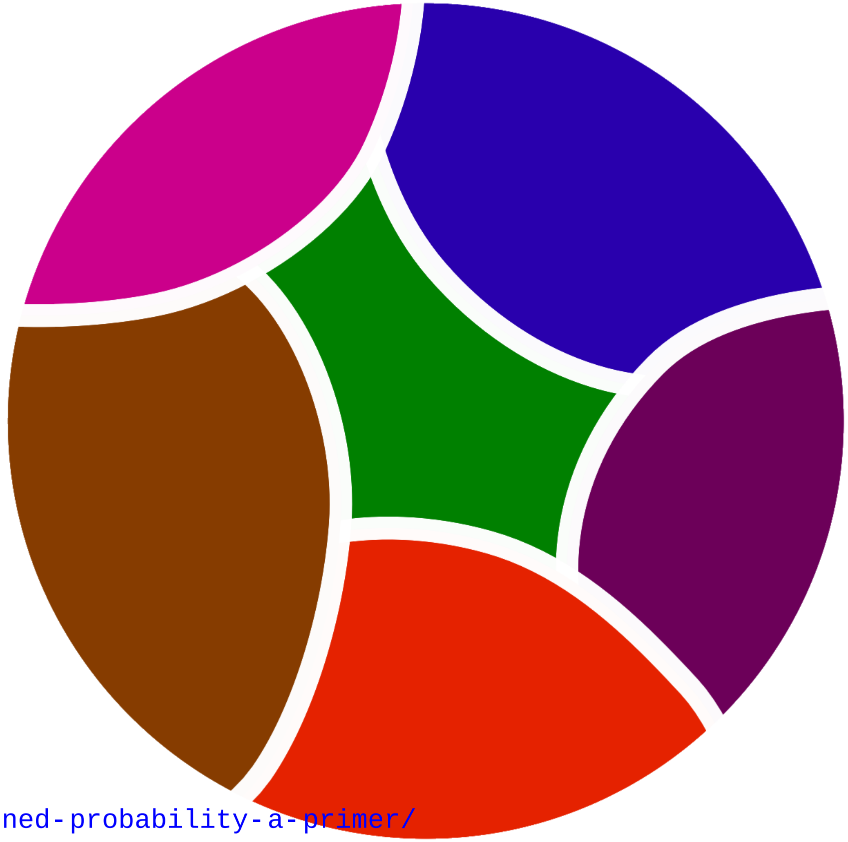
Source http://en.wikiversity.org/wiki/User:Egm6936.f10/Probability$_$concepts

Partitions and complements

De Morgans law.

$$(A \cup B)^c = A^c \cap B^c \text{ and } (A \cap B)^c = A^c \cup B^c$$

[Proof]. Venn diagrams



Partition of a set Ω. Collection of a subsets A_i such that

1. $A_i \cap A_j = \emptyset$, for $i \neq j$
2. $\cup_{i=1}^{n} A_i = \Omega$

Question. Is $\{A, A^c\}$ a partition?

Source http://jeremykun.com/2013/03/28/conditional-partitioned-probability-a-primer/

Probability

Let A be some event of the finite sample space Ω
P is a probability function if P satisfies

1. $P(A) \geq 0$
2. $P(\Omega) = 1$
3. $P(A \cup B) = P(A) + P(B)$, if A and B are disjoint

$P(A)$ is the probability that A occurs

Example. Deduce from the axioms:

- $P(A^c) = 1 - P(A)$ (use $\Omega = A \cup A^c$)
- $0 \leq P(A) \leq 1$
- $P(\emptyset) = 0$ ($\emptyset$ is the impossible event)
- $P(A \cup B) = P(A) + P(B) - P(A \cap B)$ Exercise

Probability in practice

Remark. Definition of probability does not specify how to assign probabilities to events.

Solution. Two step approach.

1. Assign probabilities to (simple) events in an experimental way:

$$\frac{n_A}{n} \longrightarrow P(A)$$

2. Use theoretical framework (definitions + derived formulas) to analyse more complicated events.

n_A: number of times event A is the outcome of an experiment; n total number of experiments

Probability example

Example.
Consider the experiment of throwing one dice twice. Let A be the event

$$A \;=\; \{\text{Total number of eyes} = 4\}$$

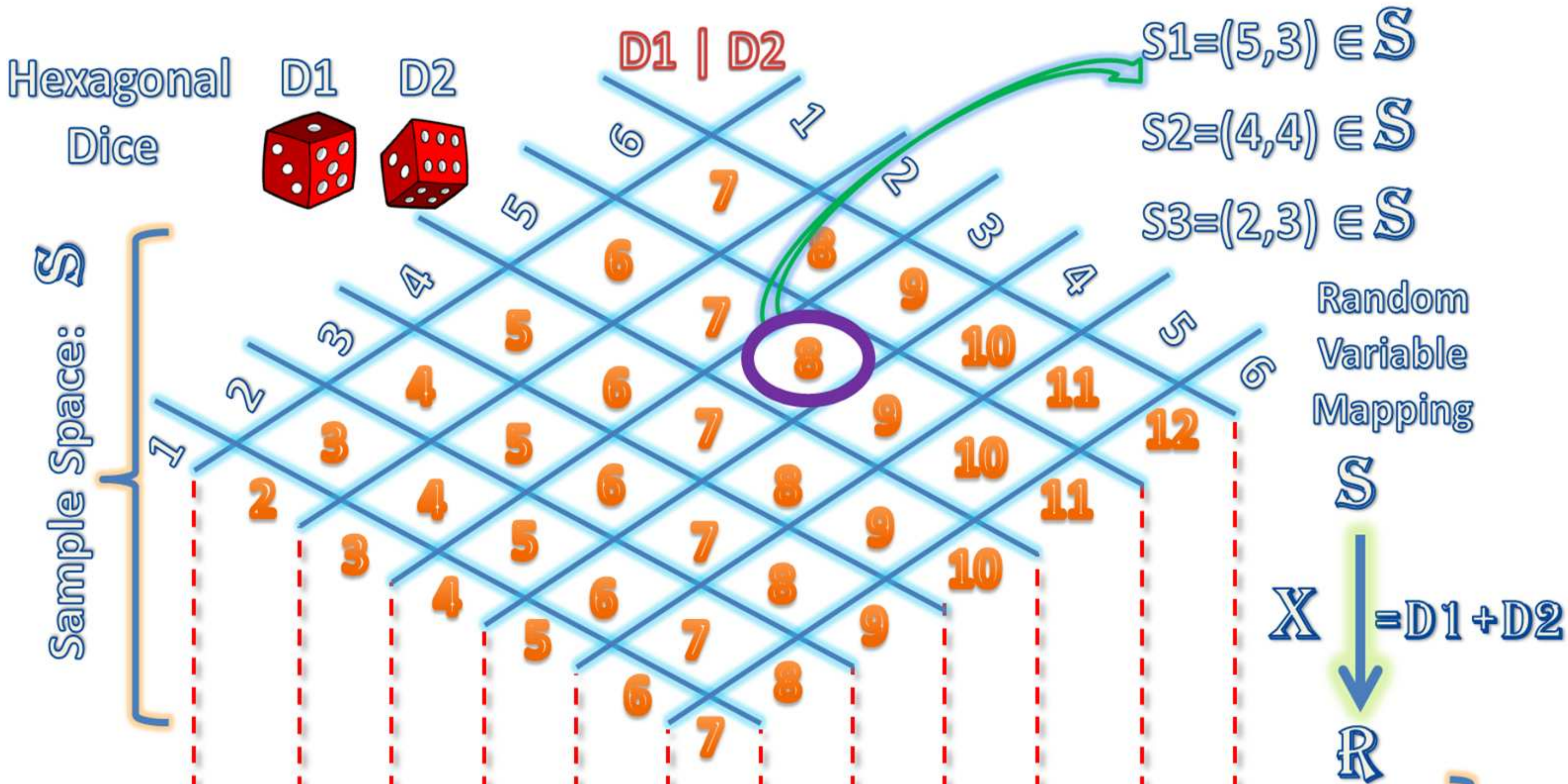
Then $P(A) = \frac{3}{36}$ as

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

while the total number of possibilities equals 36.

Question. Why 36?

Two dice probabilities



x	2	3	4	5	6	7	8	9	10	11	12	R.V.
P(X=x)	$\frac{1}{36}$	$\frac{2}{36}$	$\frac{3}{36}$	$\frac{4}{36}$	$\frac{5}{36}$	$\frac{6}{36}$	$\frac{5}{36}$	$\frac{4}{36}$	$\frac{3}{36}$	$\frac{2}{36}$	$\frac{1}{36}$	$\sum_{x=2}^{12} P(x) = 1$

Source http://wiki.stat.ucla.edu/socr/index.php/AP$_$Statistics$_$Curriculum$_$2007$_$Distrib$_$RV

Discrete sample spaces

Ω consists of n outcomes $\omega_1, \ldots, \omega_n$. An experiment is specified by the probabilities

$$p_i = P(\omega_i) \geq 0$$

of elementary events $\{\omega_i\}$ such that

$$\sum_{i=1}^{n} p_i = 1.$$

Special case. Equally likely outcomes:

$$p_1 = \cdots = p_n = \frac{1}{n}$$

Question

- Experiment where outcomes are indeed equally likely?
- Experiment where outcomes have unequal probabilities?

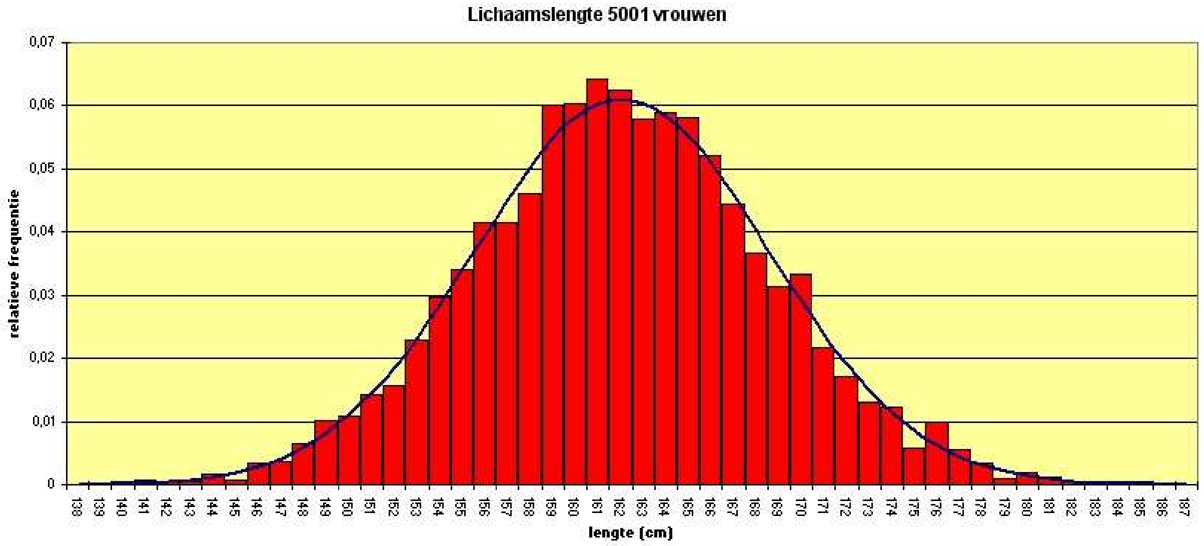
Dept. of Geoscience and Remote Sensing

Continous sample space

For example, Ω consists of all points on the real line:

$$\Omega = \mathbb{R} = (-\infty, \infty)$$

Events: all intervals $\{x_1 \leq x \leq x_2\}$.
Elementary events: $\{x_i\}$, for $x_i \in \mathbb{R}$.



Question: Is the sample space of measuring somebody's length continuous or discrete?

Probability mass

Event probabilities are specified by their probability mass:

$$P(\{x_1 \leq x \leq x_2\}) \;=\; \int_{x_1}^{x_2} \alpha(x)dx$$

where $\alpha(x)$ is a density function.

Example

$$P(\text{Dutch woman} = 1m80) \;:=\; \int_{1.795}^{1.805} \alpha(x)$$

with $\alpha(x)$ a density function describing the probability distribution of lengths of Dutch women (Like in the previous slide).

Remark. It is possible that the probability of elementary events equals zero:

$$P(\{x_i\}) \;=\; 0, \qquad\qquad \text{for all } x_i \in I\!R$$

Question. How?

B: Conditional probability

See http://math.ucsd.edu/~crypto/Monty/monty.html

Example, conditional probability

Example. Throw two dice.

A = {One of the eyes of two dice equals 1 }
What is $P(A)$?

B = {Sum of the eyes of two dice equals 4}.
What is $P(B)$?

What is $P(A)$, given that B holds? (Notation $P(A|B)$)

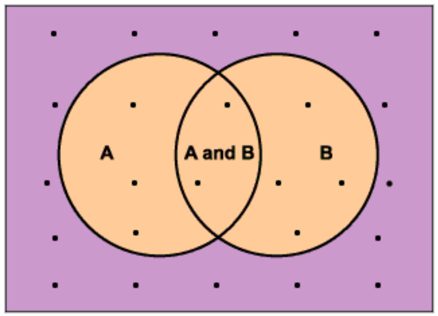
[Proof.]
$P(A)$ = first dice has a 1 + second dice has a 1, but not the first dice = $\frac{1}{6} + \frac{5}{6} \cdot \frac{1}{6} = \frac{11}{36}$

$B \ = \ \{$Sum of the eyes is $4\} \ = \ \{(1,3),(2,2),(3,1)\}$
$P(B) = \frac{3}{36}$

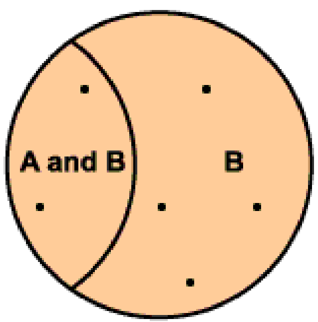
$P(A|B) = \frac{2}{3}$ (= $\frac{P(A \cap B)}{P(B)}$)

Dept. of Geoscience and Remote Sensing

Conditional probability



$\Downarrow$



Conditional probability of A given B:

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

'Reduce' sample space to the 'space' of the given outcome

That is, given B, we only care about the probability of A 'inside' B

Source http://www.quickmba.com/stats/probability/

Multiplication rule

Equivalent to the definition of conditional probability is the multiplication rule

$$P(A \cap B) \;=\; P(A|B)P(B) \;=\; (\text{symmetry}) \;=\; P(B|A)P(A)$$

Examples, conditional probability

1) $P(\Omega|B) = 1$; Why?

2) Assume A_1 and A_2 are disjoint

$$P(A_1 \cup A_2|B) \;=\; ...$$

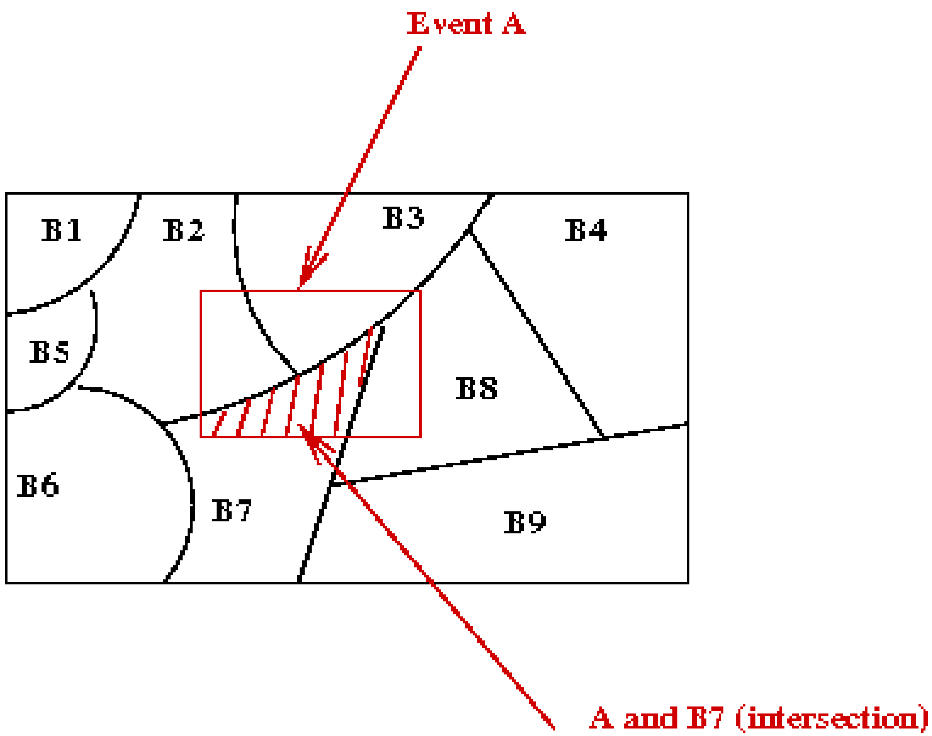
Question:
Give an example for situation 2) with double dice s.t. $A_1 \cap B \neq \emptyset$ and $A_2 \cap B \neq \emptyset$

Dept. of Geoscience and Remote Sensing

Total probability rule

Suppose that the events B_i form a partition of Ω. Then

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



[Proof.] As the B_i form a partition, it follows that

$$\begin{aligned} P(A) &= \sum_{i=1}^{n} P(A \cap B_i) \\ &= \text{(multiplication rule)} \\ &= \sum_{i=1}^{n} P(A|B_i)P(B_i) \end{aligned}$$

Source http://www.seas.gwu.edu/~simhaweb/contalg/classwork/appendix/prob/prob.html

TU Delft

Bayes rule

Assume again that the events $B_1, \ldots, B_n$ form a partition of Ω. Then, for each $j = 1, \ldots, n$:

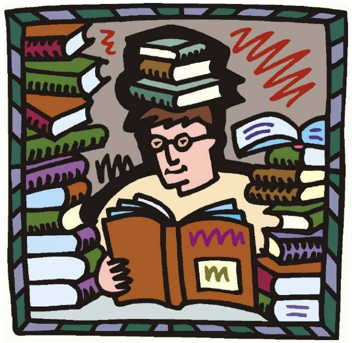
$$P(B_j|A) = \frac{P(A|B_j)P(B_j)}{\sum_{i=1}^{n} P(A|B_i)P(B_i)}$$

[Proof.] Apply the multiplication rule in two directions on $P(A \cap B_j)$. $\Rightarrow$

$$\begin{aligned} P(B_j|A)P(A) &= P(A|B_j)P(B_j), \qquad \text{or} \\ P(B_j|A) &= \frac{P(A|B_j)P(B_j)}{P(A)} \end{aligned}$$

Bayes rule follows by substituting the total probability rule for $P(A)$.

Comparing different hypotheses



At an exam a student has to answer questions by YES or NO.

His amount of study time is such that he knows the correct answer with a probability of $\frac{3}{5}$.

If he doesn't know the answer, he guesses.

1. What is the probability that an answer to an question is correct?
2. What is the probability that he indeed knew the correct answer, when his answer was correct?

Study example

Event of interest:
A = { Answer is correct }

Partition of the sample space:
B_1 = { Student knows the answer }
B_2 = { Student doesn't know the answer }

Now consider Bayes theorem:

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

The simplication follows from applying the Total probability rule as $B_1 \cup B_2 = \Omega$, i.e. B_1 and B_2 partition the sample space.

Conditional events:
$P(B_1|A)$: Probablity student knows the answer if his answer is correct,
$P(A|B_2)$: Probability answer is correct if the student doesn't know the answer, etc.

Solution to the study example

1. Probability of a correct answer:
$P(A) \; = \; P(A|B_1)P(B_1) + P(A|B_2)P(B_2) \; = \; 1 \cdot \frac{3}{5} + \frac{1}{2} \cdot \frac{2}{5} \; = \; \frac{4}{5}$

2. Probability he knew the correct answer, when his answer was correct?

$$P(B_1|A) \; = \; \frac{P(A|B_1) \cdot P(B_1)}{P(A)} \; = \; \frac{1 \cdot \frac{3}{5}}{\frac{4}{5}} \; = \; \frac{3}{4}$$

Dept. of Geoscience and Remote Sensing

Why Bayes Rule?

Bayesian spam filtering
http://en.wikipedia.org/wiki/Naive_Bayes_spam_filtering

Naive Bayes classifier
http://www.statsoft.com/textbook/naive-bayes-classifier

Independence

Two events are (mutually) independent when both

1. P(A|B) = P(A)
2. P(B|A) = P(B)

As a consequence

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

Remark. If one of these three statements is true, the other two are as well!

Example. Consider the double dice experiment. Let

- F be the event that the outcome of the first dice is 5
- G be the event that the outcome of the second dice is 5.

$P(F) = \frac{1}{6}$, $P(G) = \frac{1}{6}$; $P(F \cap G) = P(\{1, 1\}) = \frac{1}{36} = \frac{1}{6} \cdot \frac{1}{6}$, so F and G are independent.

C. Random variables and Distributions.

Dept. of Geoscience and Remote Sensing

Random variable

Let Ω be a sample space. A random variable is a function

$$X : \Omega \to I\!R$$

that assigns a real number to an event $\omega \in \Omega$.

A random variable is discrete if it takes only a finite or countable set of values.

Example.
Assume we throw a pair of dice. The function $M : \Omega \to I\!R$ that assigns the mean of the two dice to each throw is a discrete random variable.

Remark.
The concept of random variables allows us to ‘forget‘ the sample space Ω. In most cases we only need to know the probabilities of the values of the random variable X.

Dept. of Geoscience and Remote Sensing

Distribution function

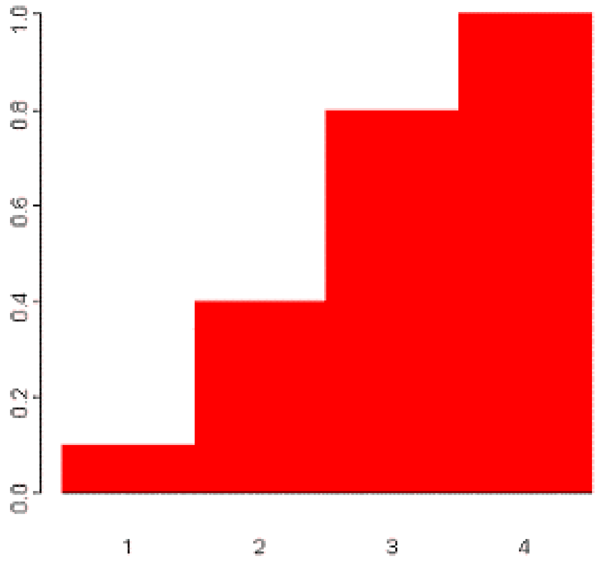
The (cumulative) distribution function F of a random variable X is the function $F : \mathbb{R} \to [0, 1]$, defined by

$$F(a) = P(X \leq a), \qquad \text{for } -\infty < a < \infty$$

Example
Suppose a random variable X can take values $1, 2, 3$ or 4. The probabilities for each outcome are as follows:

Outcome	1	2	3	4
Probability	0.1	0.3	0.4	0.2

Then X corresponds to the distribution function on the right.



Properties, distribution function

Let F be the distribution function of a random variable X. Then for F holds:

1. $a \leq b \quad \Rightarrow \quad F(a) \leq F(b)$.

   [Proof.] The event $\{X \leq a\}$ is contained in the event $\{X \leq b\}$

2. $0 \leq F(a) \leq 1$, for all $a \in \mathbb{R}$. Moreover, $F(-\infty) = 0$, and $F(\infty) = 1$.

3. $P(X > a) = 1 - F(a)$

   [Proof.] The events $\{X \leq a\}$ and $\{X > a\}$ are mutual exclusive, while their union equals Ω. Therefore, $P(X \leq a) + P(X > a) = P(\Omega) = 1$

4. $P(a < X < b) = F(b) - F(a)$

Probability mass function

Let X be a discrete random variable.

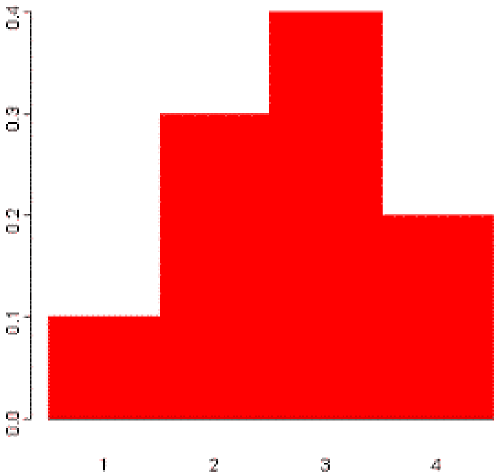
The probability mass function is a function $p : \mathbb{R} \to [0,1]$ defined as

$$p(a) \quad = \quad P(X = a), \qquad\qquad \text{for } a \in \mathbb{R}$$

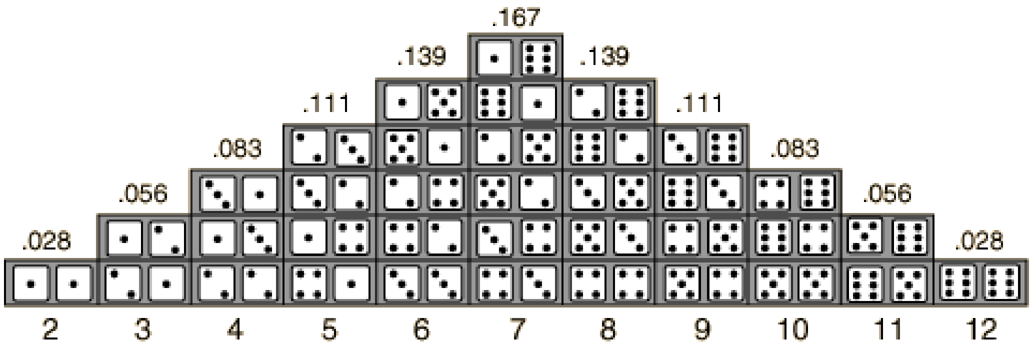
The distribution function $F(x)$ of X is a staircase function given by

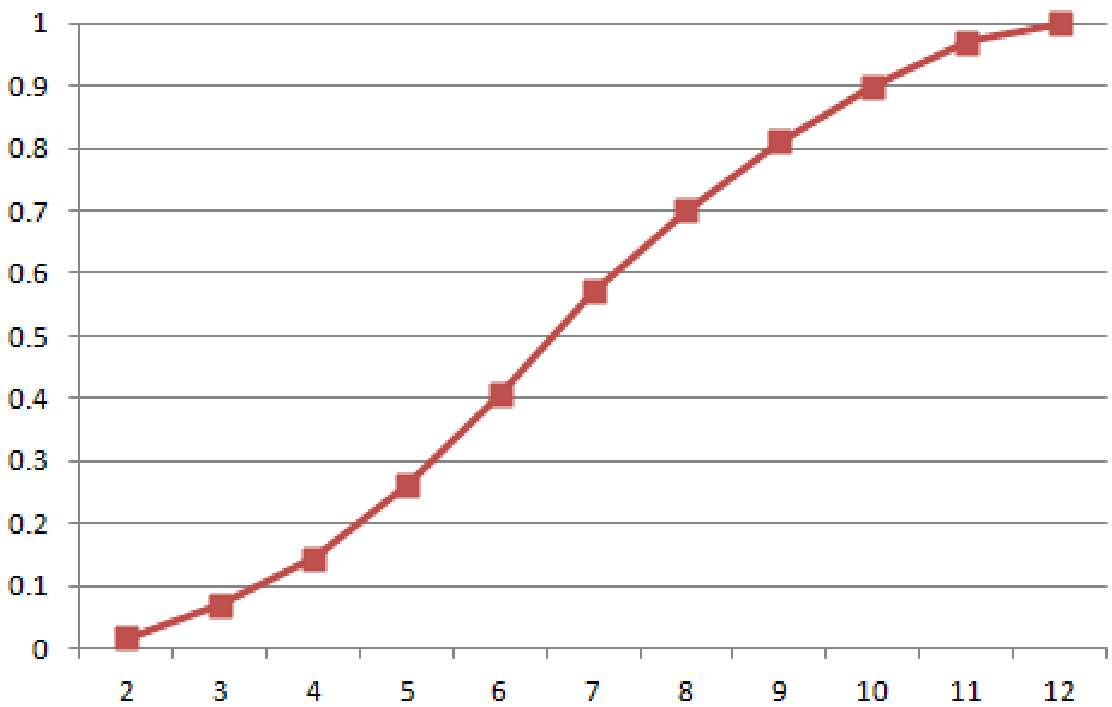
$$F(x) \quad = \quad \sum_{x_i \le x} p(x_i), \qquad\qquad \text{for } x \in \mathbb{R}$$

Example
Mass function corresponding to
Example, slide 10. $\Longrightarrow$



Mass + Distribution function, double dice





Source http://hyperphysics.phy-astr.gsu.edu/hbase/math/dice.html

Source http://stevestedman.com/2012/03/cumulative-distribution-function-cdf-analyzing-the-roll-of-dice-with-tsql/

Continuous random variable

A random variable X is continuous if for some function $f : \mathbb{R} \rightarrow \mathbb{R}$ with

1. $f(x) \geq 0$
2. $\int_{\infty}^{\infty} f(x)dx = 1$

and for any numbers a and b with $a \leq b$

$$P(a \leq X \leq b) = \int_a^b f(x)dx$$

The function f is called the probability density function of X.

Relation Density-Distribution

Let

- F be the distribution function
- f be the probability density function

of a continuous random variable. Then

$$F(b) = \int_{-\infty}^{b} f(x)dx \qquad \text{and} \qquad f(x) = \frac{d}{dx}F(x)$$

Remark: compare distribution and mass function, double dice example.

Conclusions

Probability, sample spaces, events

- Abstract, mathematical notions
- Link to real experiments is required

Conditional events:

- Dependent on other events,
- Described by e.g. Bayes theorem

Distribution, probability mass function

- Full description of all probabilities within an experiment

Exercises

Exercise 2.1 Show that for any two events:

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

Hint: write A as a combination of A and B using Venn diagrams

Exercise 2.2 Consider the experiment in which people are asked in which month they are born.

a). What is the sample space?

b). Assign a probability to each month.

c). What is the probability that three people are not born in January or February?

Exercises

Exercise 2.3 Let A and B be events in a probability space. $P(A) = 0.6$ and $P(B) = 0.3$. Further, $P(A \cup B) = 0.7$. Find the probability of the event $A^C \cup B^C$.

Exercise 2.4 We toss a fair coin three times. We define the discrete random variable X such that its value equals the number of heads in the outcome, e.g. $X(hth) = 2$.

a). What is the number of outcomes of this experiment?

b). Give the probability space.

c). What is the domain and range of X?

d). Calculate the Probability Mass Function (PMF) and make a sketch of it.

e). Calculate the Cumulative Distribution Function (CDF) and make a sketch of it.

Answers, Ex. 2.1

Show that for any two events:

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

Hint: write A as a combination of A and B using Venn diagrams

[Proof.] $A = (A \cap B) \cup (A \cap B^C)$, so also

$$P(A) = P(A \cap B) + P(A \cap B^C) \qquad (1)$$

In the same way we can write $P(A \cup B) = P((A \cup B) \cap B) \cup (A \cup B) \cap B^C))$. Note that $(A \cup B) \cap B = B$ and $(A \cup B) \cap B^C) = A \cap B^C$, So

$$P(A \cup B) = P(B) + P(A \cap B^C) \qquad (2)$$

The result follows from subtracting (2) from (1) and reordering.

Answers, Ex. 2.2

Consider the experiment in which people are asked in which month they are born.

1. What is the sample space?
2. Assign a probability to each month.
3. What is the probability that three people are not born in January or February?

[Proof.]

1. The sample space is $\{1, 2, 3, \ldots, 12\}$, where each number indicates the corresponding month.
2. To keep things simple, we may assume that each month is equally probable. In that case, $P(M_i) = 1/12$ for $i = 1, \ldots, 12$.
3. Let A be the event that a person is not born in Jan. or Feb. Then $P(A) = 10/12$. The chance that three persons are not born in Jan. or Feb. is

$$P(A)^3 = \frac{2.5.2.5.2.5}{2.2.3.2.2.3.2.2.3} = \frac{5.5.5}{6.6.6} = \frac{125}{216}$$

.

Answers, Ex. 2.3

Let A and B be events in a probability space. $P(A) = 0.6$ and $P(B) = 0.3$. Further, $P(A \cup B) = 0.7$. Find the probability of the event $A^C \cup B^C$.

[Proof.] $P(A^C \cup B^C) \ = \ 1 - P(A \cap B)$.
Moreover, $P(A \cap B) \ = \ P(A) + P(B) - P(A \cup B) \ = \ 0.6 + 0.3 - 0.7 \ = \ 0.2$
$\Rightarrow P(A^C \cup B^C) \ = \ 0.8$

Answers, Ex. 2.4

We toss a fair coin three times. We define the discrete random variable X such that its value equals the number of heads in the outcome, e.g. $X(hth) = 2$.

1. What is the number of outcomes of this experiment?
2. Give the probability/sample space.
3. What is the domain and range of X?
4. Calculate the Probability Mass Function (PMF) and make a sketch of it.
5. Calculate the Cumulative Distribution Function (CDF) and make a sketch of it.

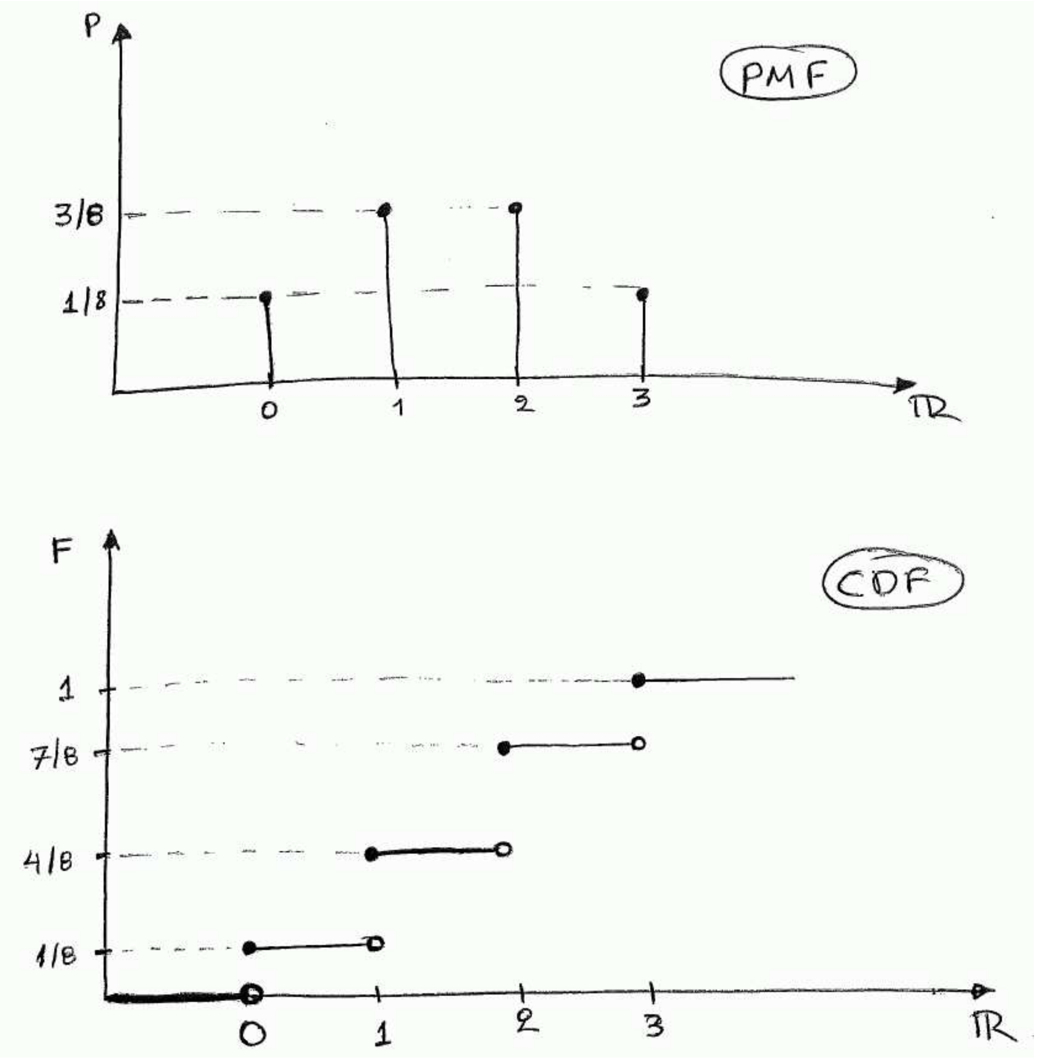
[Proof.]

1. Possible outcomes: 0, 1, 2, 3.
2. $\{ttt\} \to 0$, $\{tht, htt, tth\} \to 1$, $\{hth, hht, thh\} \to 2$, $\{hhh\} \to 3$.
   $P(X = 0) = \frac{1}{8}$; $P(X = 1) = \frac{3}{8}$; $P(X = 2) = \frac{3}{8}$; $P(X = 3) = \frac{1}{8}$.
3. Domain: $\{ttt, tht, htt, tth, hth, hht, thh, hhh\}$. Range: {0,1,2,3}.

Dept. of Geoscience and Remote Sensing

TU Delft

Answers, Ex. 2.4, iv and v



Dept. of Geoscience and Remote Sensing

Links for the next lectures

Source http://www.intechopen.com/source/html/43533/media/image1.png

Source http://en.wikiversity.org/wiki/User:Egm6936.f10/Probability$_$concepts

Source http://wiki.stat.ucla.edu/socr/index.php/AP$_$Statistics$_$Curriculum$_$2007$_$Distrib$_$RV

Source http://www.math.uah.edu/stat/prob/Probability.html

Source http://en.wikibooks.org/wiki/Probability/Print$_$version

Source http://www.colorado.edu/geography/gcraft/notes/error/error$_$f.html