Probability Summary

Ánoq of the Sun, Hardcore Processing *

July 1, 2004

1 Probability

1.1 Intro, Notation, Terminology

Premises: Ω ***sample space***, $\mathcal{A}$ ***all possibilities of*** Ω

- **The *sample space* Ω $\langle greek : \delta.\chi.\rangle$ $\stackrel{def}{=}$** $_{\langle \text{p. } 21 \in [1]\rangle}$
  The *set* of *all* the *possible results* of an *experiment of chance* $\langle greek : \pi.\tau.\rangle$.
  Ω may be *discrete* (i.e. *finite* or *countable*) or *continuous*, i.e.:
  $\Omega = \{\omega_1, \ldots, \omega_n\}$: *finite* (*discrete*)
  $\Omega = \{\omega_1, \omega_2, \ldots\}$: *countably infinite* (*discrete*)
  $\Omega = \{\omega \in \mathbb{R} \mid 0 \leq \omega \leq 1\}$: *Over-countably infinite* (*continuous*)

- **For $\omega \in \Omega$, ω is called a *sample point* / *basic possibility*** $_{\langle \text{p. } 21 \in [1]\rangle}$

- **The *set of all possibilities* $\mathcal{A}$ of Ω $\stackrel{def}{=}$** $_{\langle \text{def. 2.1 p. } 22 \in [1]\rangle}$
  Either the *power set* of Ω: $\mathcal{A} = \mathcal{P}(\Omega)$ (always the case for *discrete* Ω)
  or a *subset* of the *power set*: $\mathcal{A} \subset \mathcal{P}(\Omega)$ (only for *continuous* Ω)
  In general $\mathcal{A}$ *must also* be a *σ-field* or a *σ-algebra*, which means:

  1. $\Omega \in \mathcal{A}$
  2. $(A \in \mathcal{A}) \Rightarrow (\complement A \in \mathcal{A})$
  3. $(\forall \nu \in J : A_\nu \in \mathcal{A}) \Rightarrow (\cap_{\nu \in J} A_\nu \in \mathcal{A})$ for *countable index set* J

  From these it follows that e.g. also *countable unions* and $\emptyset$ are in $\mathcal{A}$.

- **For $A \subset \Omega$, A is called a *possible outcome*** $_{\langle \text{def. 2.1 p. } 22 \in [1]\rangle}$
  E.g.: For a *die roll* of a *6-sided die*: $\Omega = \{1, \ldots, 6\}$, $\omega = 4$, $A = \{1, 2, 3\}$

- **For $A \in \mathcal{A}$, A is a *possible outcome*** $_{\langle \text{def. 2.1 p. } 22 \in [1]\rangle}$

1.2 Different Probability Theories

- *Theory of Laplace* (for *discrete* Ω):
  If all *possible outcomes* have the same probability.
  Assume that a *possible outcome* A results from $N(A)$ elements of Ω, then $P(A) = \frac{N(A)}{N}$, when $N = N(\Omega)$. ⟨p. 26 ∈ [1]⟩

- *Geometric Probabilities*: Let Ω be an *over-countable set* with *measure* $\mu(\Omega)$ (*area*, *volume*, etc.) and let $A \subset \Omega$ with $\mu(A)$ being the measure.
  If all *possibilities* of Ω have the *same probabilities*, then $P(A) = \frac{\mu(A)}{\mu(\Omega)}$ ⟨p. 27 ∈ [1]⟩

- *Theory of Von Mises*:
  If the *relative frequency* of the *possible outcome* A is $R_\nu(A) = \frac{m_\nu(A)}{\nu}$, then $P(A) = \lim_{\nu\to\infty} R_\nu(A)$. ⟨p. 40 ∈ [1]⟩

- For both Laplace and Von Mises Theory the *axioms of Kolmogorov* holds with *finite additivity* instead of σ*-additivity.* ⟨p. 26 ∈ [1]⟩

- ***Homeomorphic probabilities*** $\stackrel{def}{=}$
  $\forall j \in \{1, \dots, N\} : P(\{\omega_j\}) = \frac{1}{N}$ ⟨rem 6.1 (i) p. 42 ∈ [1]⟩

1.3 Counting the Number of Elements $N(A)$ of a Possible Outcome A

- For a *finite sample space* Ω with ν *elements*: The *power set* $\mathcal{P}(\Omega)$ is:
  $\binom{\nu}{0} + \binom{\nu}{1} + \cdots + \binom{\nu}{\nu} = (1+1)^\nu = 2^\nu$

- For *finite disjoint* $A_1, A_2, \dots, A_n$:
  $N(A_1 \cup A_2 \cup \cdots \cup A_n) = N(A_1) + N(A_2) + \cdots + N(A_n)$
  (*summation principle* of combinatorics) ⟨p. 29 ∈ [1]⟩

- For *finite disjoint* $A_1, A_2, \dots, A_n$:
  $N(A_1 \times A_2 \times \cdots \times A_n) = N(A_1)N(A_2)\cdots N(A_n)$
  (*multiplication principle* of combinatorics) ⟨p. 29 ∈ [1]⟩

1.4 The Axiomatic Foundation by Kolmogorov

Let Ω be the *sample space* of an *experiment of randomness* with $\mathcal{A}$ being the *σ-algebra of possibilities.*
Probability is a *set function* $P : \mathcal{A} \to \mathbb{R}$ which *complies with* the *following*:

- Axioms: ⟨def. 6.1 p. 41 ∈ [1]⟩

  1. $\forall A \in \mathcal{A} : P(A) \geq 0$
  2. $P(\Omega) = 1$
  3. $P(\cup_{i=1}^{\infty} A_i) = \sum_{i=1}^{\infty} P(A_i)$ for *pairwise disjoint* $A_i \in \mathcal{A}$. (*σ-additivity*)

  The probability space $\stackrel{def}{=} (\Omega, \mathcal{A}, P)$

- Theorems (from the axioms), where $A, B \in \mathcal{A}$:

  1. $P(\emptyset) = 0$ ⟨thm. 6.1 (i) p. 43 ∈ [1]⟩
  2. $\sum_{i=1}^{\nu} P(A_i) = P(\cup_{i=1}^{\nu} A_i)$ for *pairwise disjoint* $A_i \in \mathcal{A}$ (*finite additivity*) ⟨thm. 6.1 (ii) p. 43 ∈ [1]⟩
  3. $P(\complement A) = 1 - P(A)$ ⟨thm. 6.1 (iii) p. 43 ∈ [1]⟩
  4. $P(A \backslash B) = P(A) - P(A \cap B)$
     And if $B \subset A$, then $P(A) - P(B)$ ⟨thm. 6.1 (iv) p. 43 ∈ [1]⟩
  5. $0 \leq P(A) \leq 1$ ⟨thm. 6.2 p. 45 ∈ [1]⟩
  6. For $A \subset B : P(A) \leq P(B)$ - i.e. P is a *growing set function* ⟨thm. 6.2 p. 45 ∈ [1]⟩
  7. For $A_i \in \mathcal{A}$: $P(\cup_{i=1}^{\infty} A_i) \leq \sum_{i=1}^{\infty} P(A_i)$ ⟨thm. 6.3 p. 48 ∈ [1]⟩
     (*Boole's inequality* / *σ-subadditivity* / *countable subadditivty*)

- Theorems similar to the *inclusion principle* (and it's inversion), $A, B \in \mathcal{A}$:

  1. $P(A \cup B) = P(A) + P(B) - P(A \cap B)$ ⟨thm. 7.1p. 53 ∈ [1]⟩
  2. $P(\complement A \cap \complement B) = 1 - P(A) - P(B) + P(A \cap B)$ ⟨cor. 7.1p. 53 ∈ [1]⟩
  3. Poincarè's Theorem: For $A_1, \ldots, A_\nu \in \mathcal{A}$: ⟨thm. 7.2 p. 54 ∈ [1]⟩
     $P(A_1 \cup \cdots \cup A_\nu) = S_{\nu,1} - S_{\nu,2} + \cdots + (-1)^{\nu-1} S_{\nu,\nu}$, where
     $S_{\nu,1} = \sum_{i=1}^{\nu} P(A_i)$,
     $S_{\nu,2} = \sum_{i=1}^{\nu-1} \sum_{j=i+1}^{\nu} P(A_i \cap A_j)$,
     $\vdots$
     $S_{\nu,r} = \sum P(A_{i_1} \cap A_{i_2} \cap \cdots \cap A_{i_r})$, where the *sum* comes from *all the coices* of the r *indices* $\{i_1, \ldots, i_r\}$ *out of* $\{1, \ldots, \nu\}$,
     $\vdots$
     $S_{\nu,\nu} = P(A_1 \cap A_2 \cap \cdots \cap A_\nu)$
  4. For $A_1, \ldots, A_\nu \in \mathcal{A}$: ⟨cor. 7.2 p. 54 ∈ [1]⟩
     $P(\complement A_1 \cap \cdots \cap \complement A_\nu) = 1 - S_{\nu,1} + S_{\nu,2} + \cdots + (-1)^{\nu} S_{\nu,\nu}$, where $S_{q,r}$ are as in *Poincarè's theorem.*

1.5 Special Cases

- For $A_1, \ldots, A_\nu \in \mathcal{A}$: **The A_i are *exchangeable*** $\stackrel{def}{=}$ ⟨def. 7.1 p.57 ∈ [1]⟩
  $\forall r \in \{1, \ldots, \nu\} : P(A_{i_1} \cap A_{i_2} \cap \cdots \cap A_{i_r}) = P_r$ for
  any choice of *indices* $\{i_1, \ldots, i_r\}$.
  I.e.: $P(A_1) = P(A_2) = \cdots = P(A_\nu)$,
  $P(A_1 \cap A_2) = P(A_1 \cap A_3) = \cdots = P(A_{\nu-1} \cap A_\nu)$,
  $P(A_1 \cap A_2 \cap \cdots \cap A_r) = P(A_2 \cap A_3 \cap \cdots \cap A_{r+1}) = \cdots$

- For *exchangeable* $A_1, \ldots, A_\nu \in \mathcal{A}$: ⟨cor. 7.3 p.57 ∈ [1]⟩

  1. $P(A_1 \cup \cdots \cup A_\nu) = \binom{\nu}{1} P(A_1) - \binom{\nu}{2} P(A_1 \cap A_2) + \cdots +$
     $(-1)^{r-1} \binom{\nu}{r} P(A_1 \cap \cdots \cap A_r) + \cdots + (-1)^{\nu-1} \binom{\nu}{\nu} P(A_1 \cap \cdots \cap A_\nu)$
  2. $P(\complement A_1 \cap \cdots \cap \complement A_\nu) =$
     $1 - \binom{\nu}{1} P(A_1) + \cdots + (-1)^\nu \binom{\nu}{\nu} P(A_1 \cap \cdots \cap A_\nu)$

1.6 Bound Probabilities (A Posteriori Probabilities)

Premises: $(\Omega, \mathcal{A}, P)$ ***probability space***, $A, B \in \mathcal{A}$, $P(A) > 0$ c

- **The *probability* of B *bound by* A** (when $P(A) > 0$) $\stackrel{def}{=}$
  the *set function* $P(B|A) = \frac{P(A \cap B)}{P(A)} = \frac{N(A \cap B)}{N(A)}$ ⟨def. 1.1 p. 80 ∈ [1]⟩

- The *axioms of Kolmogorov* also hold for $P(B|A)$: ⟨def. 1.1 p. 80 ∈ [1]⟩

  1. $P(B|A) \geq 0$
  2. $P(\Omega|A) = 1$
  3. $P(\cup_{i=1}^\infty B_i|A) = \sum_{i=1}^\infty P(B_i|A)$ for *pairwise disjoint* $B_i \in \mathcal{A}$

  Thus $P(\cdot|A)$ is a *probability function*, so the usual theorems apply. e.g: ⟨rem. 1.1. p. 80 ∈ [1]⟩

  1. $P(\complement B|A) = 1 - P(B|A)$
     Warning: $P(B|\complement A) = 1 - P(B|A)$ does *not hold*!
  2. $P(B_1 \cup B_2|A) = \frac{P(B_1 \cap A) + P(B_2 \cap A) - P(B_1 \cap B_2 \cap A)}{P(A)}$

- Multiplication theorem: For $\forall i \in \{1, \ldots, \nu\} : A_i \in \mathcal{A}$ and
  $P(A_1 \cap A_2 \cap \cdots \cap A_{\nu-1}) > 0$: ⟨thm. 1.1 p. 81 ∈ [1]⟩
  $P(A_1 \cap A_2 \cap \cdots \cap A_\nu) =$
  $P(A_1) P(A_2|A_1) P(A_3|A_1 \cap A_2) \cdots P(A_\nu|A_1 \cap \cdots \cap A_{\nu-1})$

- Theorem of all probabilities: For *partition* $\{A_i \mid i \in \mathbb{N}\}$ of *sample space* Ω
  where $\forall i \in \mathbb{N} : P(A_i) > 0$ we have:
  $\forall B \in \mathcal{A} : P(B) = \sum_{i=1}^\infty P(A_i) P(B|A_i)$. ⟨thm. 1.2 p. 84 ∈ [1]⟩

- Bayes theorem: For *partition* $\{A_i \mid i \in \mathbb{N}\}$ of *sample space* Ω where
  $\forall i \in \mathbb{N} : P(A_i) > 0$ and a *possible outcome* $B \in \mathcal{A}$ where $P(B) > 0$:
  $\forall i \in \mathbb{N} : P(A_i|B) = \frac{P(A_i) P(B|A_i)}{\sum_{k=1}^\infty P(A_k) P(B|A_k)}$. ⟨thm. 1.3 p. 84 ∈ [1]⟩

- *The theorem of all probabilities* and *Bayes theorem* also hold for
  finite partitions. ⟨rem. 1.3 p. 85 ∈ [1]⟩

1.7 Stochastic Independence

Premises: $(\Omega, \mathcal{A}, P)$ ***probability space***, $A, B \in \mathcal{A}$

- For A, B *disjoint*: $P(B|A) = 0$ ⟨p. 91 (i) ∈ [1]⟩
- For $A \subset B$ (so: $A \cap B = B$): $P(B|A) = 1$ ⟨p. 91 (ii) ∈ [1]⟩
- For $B \subsetneq A \subsetneq \Omega$ (so: $A \cap B = B$): $P(B|A) > P(B)$ ⟨p. 91 (iii) ∈ [1]⟩
- For $\complement B \subsetneq A \subsetneq \Omega$ (so: $A \cap \complement B = \complement B$): $P(B|A) < P(B)$ ⟨p. 91 (iv) ∈ [1]⟩
- For $P(A), P(B) > 0$: $P(A \cap B) = P(A)P(B|A) = P(B)P(A|B)$ ⟨p. 91 ∈ [1]⟩
- B **is *stochasticly independent* of** $A \overset{def}{=} P(B|A) = P(B)$ ⟨p. 91 (v) ∈ [1]⟩
- If B is *stochasticly independent* of A (or vice versa) then:
  $P(A \cap B) = P(A)P(B)$ ⟨p. 91 ∈ [1]⟩
- A **and** B **are *stochasticly independent*** $\overset{def}{=}$
  $P(A \cap B) = P(A)P(B)$ ⟨def. 2.1 p. 92 ∈ [1]⟩
- If A and B are *stochasticly independent*, then so is:
  A and $\complement B$, $\complement A$ and B, $\complement A$ and $\complement B$. ⟨p. 93 ∈ [1]⟩

2 Combinatorics

- *Summation Principle*: Assume that an *element* α_i can be *chosen* in κ_i *different ways*, where *each of those precludes simultaneous choice* of *another element* α_j, $j \in \{1, 2, 3, \ldots, \nu\}, i \neq j$.
Then the *elements*: α_1 *or* α_2 *or* $\cdots$ *or* α_ν can be *chosen* in:
$\kappa_1 + \cdots + \kappa_\nu$ *different ways.* ⟨p. 29 ∈ [1]⟩

- *Multiplication Principle*: Assume that an *element* α_1 can be *chosen* in κ_1 *different ways* and for *each of those* there are κ_2 *different ways* to choose α_2 etc.
Then the *elements*: α_1 *and* α_2 *and* $\cdots$ *and* α_ν can be *chosen* in:
$\kappa_1 \kappa_2 \cdots \kappa_\nu$ *different ways.* ⟨p. 29 ∈ [1]⟩

- Assume that we have ν *distinct elements* $(\alpha_1, \alpha_2, \ldots, \alpha_\nu)$, then:
Possible orders: $\nu \cdot (\nu - 1) \cdot (\nu - 2) \cdots 2 \cdot 1 = \nu!$ ⟨p. 30 ∈ [1]⟩
Orders (permutations) for κ out of ν elements: $\nu \cdot (\nu - 1) \cdots (\nu - \kappa + 1) = \frac{\nu!}{(\nu-\kappa)!} = (\nu)_\kappa$ ⟨thm. 4.1α p. 30 ∈ [1]⟩
Orders for κ out of ν elements with repetition: $U(\nu, \kappa) = \nu \cdot \nu \cdots \nu = \nu^\kappa$ ⟨thm. 4.1β p. 30 ∈ [1]⟩

- **Binomial Coefficient**: *Combinations of κ out of ν elements*: ⟨thm. 4.1β p. 30 ∈ [1]⟩
$\frac{(\nu)_\kappa}{\kappa!} = \frac{\nu \cdot (\nu-1) \cdots (\nu-\kappa+1)}{\kappa!} = \frac{\nu \cdot (\nu-1) \cdots 2 \cdot 1}{\kappa! \cdot 1 \cdot 2 \cdots (\nu-\kappa)} = \frac{\nu!}{\kappa!(\nu-\kappa)!} = \binom{\nu}{\kappa}$

- ***Combinations* of κ *out of* ν *elements* with *repetition***: ⟨thm. 4.2β p. 31 ∈ [1]⟩
$\left[\begin{array}{c} \nu \\ \kappa \end{array}\right] = \binom{\nu - 1 + \kappa}{\kappa} = \binom{\nu - 1 + \kappa}{\nu - 1}$

- If we place the *elements* $\alpha_1, \ldots, \alpha_\nu$ into r *ordered subpopulations* with $\kappa_1, \kappa_2, \ldots, \kappa_r$ *elements* respectively, where $\kappa_1 + \kappa_2 + \cdots + \kappa_r = \nu$ then:

  1. There are $C(\nu, \kappa_1, \kappa_2, \ldots, \kappa_{r-1}) = \binom{\nu}{\kappa_1, \kappa_2, \ldots, \kappa_{r-1}} = \frac{\nu!}{\kappa_1! \kappa_2! \cdots \kappa_r!}$ *possibilities.* ⟨thm. 4.3α p. 32 ∈ [1]⟩
  2. $C(\nu, \kappa_1, \kappa_2, \ldots, \kappa_{r-1}) = \binom{\nu}{\kappa_1} \cdot \binom{\nu - \kappa_1}{\kappa_2} \cdots \binom{\nu - (\kappa_1 + \cdots + \kappa_{r-2})}{\kappa_{r-1}}$ ⟨thm. 4.3 proof p. 32 ∈ [1]⟩
  3. We have ν *elements* of which κ_1 are *alike*, κ_2 are *alike*, ..., κ_r are *alike*. There are $M(\kappa_1, \kappa_2, \ldots, \kappa_r) = C(\nu, \kappa_1, \kappa_2, \ldots, \kappa_{r-1}) = \frac{\nu!}{\kappa_1! \kappa_2! \cdots \kappa_r!}$ *permutations.* ⟨thm. 4.3β p. 32 ∈ [1]⟩

- *Binomial Development*:
$(\alpha + \beta)^\nu = \sum_{\kappa=0}^{\nu} \binom{\nu}{\kappa} \alpha^\kappa \beta^{\nu-\kappa}$

- *Relations of Pascal's Triangle*:
$\binom{\nu}{\kappa} = \binom{\nu}{\nu - \kappa} = \binom{\nu - 1}{\kappa - 1} + \binom{\nu - 1}{\kappa}$

- *Stirling's Formula*: $\nu! \approxeq e^{-\nu} \nu^{\nu + \frac{1}{2}} \sqrt{2\pi}$ ⟨thm. 4.4 p. 33 ∈ [1]⟩

2.1 Combinatorical Principles

- Putting κ *distinct speres* into ν *distinct cells* with *unlimited space*:
  ν^{κ} *ways.* ⟨ex. 4.1α p. 36 $\in$ [1]⟩

- Putting κ *distinct speres* into ν *distinct cells* where
  cell j contains κ_j spheres and $\kappa_1 + \kappa_2 + \cdots + \kappa_\nu = \kappa$:
  $C(\nu, \kappa_1, \kappa_2, \ldots, \kappa_{r-1}) = \frac{\kappa!}{\kappa_1! \kappa_2! \cdots \kappa_r!}$ *ways.* ⟨ex. 4.1α p. 37 $\in$ [1]⟩

- Putting κ *identical speres* into ν *distinct cells* corresponds to
  choosing κ out of ν cells: ⟨ex. 4.1β p. 37 $\in$ [1]⟩

  - If each cell can *contain 1 sphere*, there are:
    $\begin{pmatrix} \nu \\ \kappa \end{pmatrix}$ *ways.*
  - If each cell has *unlimited space*, there are:
    $\begin{bmatrix} \nu \\ \kappa \end{bmatrix} = \begin{pmatrix} \nu + \kappa - 1 \\ \kappa \end{pmatrix}$ *ways.*

- Ways of choosing κ *ordered elements without repetition* out of ν elements:
  $(\nu)_\kappa$ ⟨ex. 4.2α p. 28 $\in$ [1]⟩

- Ways of choosing κ *unordered elements without repetition* out of ν el.:
  $\begin{pmatrix} \nu \\ \kappa \end{pmatrix}$ ⟨ex. 4.2β p. 28 $\in$ [1]⟩

- Ways of choosing κ *ordered elements with repetition* out of ν elements:
  ν^{κ} ⟨ex. 4.2γ p. 28 $\in$ [1]⟩

- Ways of choosing κ *unordered elements with repetition* out of ν elements:
  $\begin{bmatrix} \nu \\ \kappa \end{bmatrix} = \begin{pmatrix} \nu + \kappa - 1 \\ \kappa \end{pmatrix}$ ⟨ex. 4.2δ p. 28 $\in$ [1]⟩

3 Set Theorems

The following holds for *any algebra* and for the usual *set theory*.

- For $A, B, \Gamma \in \mathcal{A}$ we have: ⟨p. 5-10 ∈ [1]⟩

Associativity:	$(A \cup B) \cup \Gamma = A \cup (B \cup \Gamma)$
	$(A \cap B) \cap \Gamma = A \cap (B \cap \Gamma)$
Distributivity:	$A \cap (B \cup \Gamma) = (A \cap B) \cup (A \cap \Gamma)$
	$A \cup (B \cap \Gamma) = (A \cup B) \cap (A \cup \Gamma)$
Commutativity:	$A \cup B = B \cup A$
	$A \cap B = B \cap A$
Misc:	$A \cup \emptyset = A$
	$A \cap \Omega = A$
	$A \cup \Omega = \Omega$
	$A \cap \emptyset = \emptyset$
	$\complement \Omega = \emptyset$
	$\complement \emptyset = \Omega$
	$\complement\complement A = A$
	$A \cup \complement A = \Omega$
	$A \cap \complement A = \emptyset$
De Morgan:	$\complement(\cup_{\nu=1}^{\kappa} A_\nu) = \cap_{\nu=1}^{\kappa}(\complement A_\nu)$ (also for infinitely many on a σ-algebra)
	$\complement(\cap_{\nu=1}^{\kappa} A_\nu) = \cup_{\nu=1}^{\kappa}(\complement A_\nu)$ (also for infinitely many on a σ-algebra)

4 Misc

- **A *partition* of a *sample space*** $\Omega \overset{def}{=}$ a *sequence* of *possible outcomes* $\forall i \in \mathbb{N} : A_i \in \mathcal{A}$ which are *pairwise disjoint* (i.e. $\forall i \neq j : A_i \cap A_j = \emptyset$) and where $\cup_{i=1}^{\infty} A_i = \Omega$. ⟨p. 84 ∈ [1]⟩

References

[1] Χαραλαμπού Α. Χαραλαμπίδη. *Θεωρία Πιθανοτήτων και Εφαρμογές*, Εκδόσεις Συμμετρία Αθήνα 2000.