Probability
Exercises

Jacek Cichoń
WPPT, PWr, Big Data Analytics, 2018/19

1 Probability spaces

Exercise 1 — Let $(A_k)_{k=1\ldots n}$ be a sequence of subsets of a space Ω. For $\varepsilon \in \{0,1\}^n$ we put

$$A_\varepsilon = \bigcap_{k=1}^{n} A_k^{\varepsilon_k}$$

where $A^0 = \Omega \setminus A$ and $A^1 = A$.

1. Show that if $\varepsilon, \eta \in \{0,1\}^n$ and $\varepsilon \neq \eta$ then $A_\varepsilon \cap A_\eta = \emptyset$.
2. Show that $\bigcup\{A_\varepsilon : \varepsilon \in \{0,1\}^n\} = \Omega$.
3. Conclude that $\{A_\varepsilon : \varepsilon \in \{0,1\}^n\}$ is a partition of the space Ω.
4. For $x, y \in \Omega$ we define $(x \equiv y) \leftrightarrow (\forall k = 1\ldots n)(x \in A_k \leftrightarrow y \in A_k)$. Show that $\equiv$ is an equivalence relation on the set Ω and that
$$(\Omega/\equiv) = \{A_\varepsilon : \varepsilon \in \{0,1\}^n\}\ .$$

Exercise 2 — Let $\Omega = \{1, 2, \ldots, 10\}$. Let $A = \{2, 3, 4\}$ and $B = \{3, 4, 5, 6\}$.

1. Describe the σ-field of subsets of Ω generated by the family $\{A, B\}$.
2. What is the cardinality of this σ-field ?
3. Show an example of subset of Ω which is not in this σ-field.

Exercise 3 — Let $\Omega = \{1, 2, \ldots, 10\}$. Let $A = \{3, 4\}$ and $B = \{3, 4, 5, 6\}$.

1. Describe the σ-field of subsets of Ω generated by the family $\{A, B\}$.
2. What is the cardinality of this σ-field ?

Exercise 4 — Let $\Omega = [0, 1]$. Let $\mathcal{A}$ be σ-field of subsets of Ω. Let $\mathcal{B} = \{A \times \Omega : A \in \mathcal{A}\}$.

1. Show that $\mathcal{B}$ is a σ-field of subsets of $[0, 1]^2$.
2. Show examples of subsets of $[0, 1]^2$ which are not in $\mathcal{B}$.

Exercise 5 — Suppose that $A, B \in \mathrm{Bor}([0, 1])$. Show that $A \times B \in \mathrm{Bor}([0, 1]^2)$.
Hint: Observe that $A \times B = (A \times [0, 1]) \cap ([0, 1] \times B)$.

Exercise 6 — Suppose that $f : [0, 1] \to [0, 1]$ is continuous.

1. Show that the graph of f, i.e. the set $\{(x, f(x)) : x \in [0, 1]\}$, is a Borel subset of $[0, 1]^2$.
2. Show that the set $\{(x, y) \in [0, 1]^2 : y \leq f(x)\}$ is a Borel subset of $[0, 1]^2$.

1.1 Discrete probability spaces

Exercise 7 — Let $\Omega = \{0, 1, \ldots, 100\}$. We consider combinatorial (counting) probabilities. Calculate probabilities of the following events:

1. $\{(x, y) \in \Omega^2 : x = y\})$
2. $\{(x, y) \in \Omega^2 : x < y\})$

3. $\{(x, y) \in \Omega^2 : x + y \leq 50\})$

Exercise 8 — Show that if $|A| = n$, then $|P(A)| = 2^n$.

Exercise 9 — Let us fix two numbers $n, k \in \mathbb{N}$. Let $[n]^k = \{X \subseteq \{1, \ldots, n\} : |X| = k\}$. Show that $|[n]^k| = \binom{n}{k}$.

Exercise 10 — We choose from the set $\{1, \ldots, 10\}$ according to uniform distribution a subset of cardinality 5. We model this experiment as follows: we consider the combinatorial probability space $\Omega = \{A \subseteq \{1, \ldots, 10\} : |A| = 5\}$. Calculate the probability of the event „the middle element of this set jest 5"? In othe words: calculate the probability of the event

$$A = \{X \in \Omega : 5 \in X \wedge |X \cap \{1, 2, 3, 4\}| = |\{X \cap \{6, 7, 8, 9, 10\}| = 2\} .$$

Exercise 11 — Let $\Omega = \mathbb{N}$, $\mathcal{S} = P(\Omega)$ and $\Pr(A) = \sum_{n \in A} \frac{1}{2^{n+1}}$ for $A \subseteq \mathbb{N}$.

1. Check that Pr is a probability.
2. Calculate $\Pr(\{n \in \mathbb{N} : n \geq 5\})$ oraz $\Pr(\{2n : n \in \mathbb{N}\})$.

Exercise 12 — Let $\mathcal{P} = (\Omega, \mu)$ be a finite probability space. We define a function $\lambda : P(\Omega^2) \to \mathbb{R}$ by formula

$$\lambda(A) = \sum \{\mu(x) \cdot \mu(y) : (x, y) \in A\} .$$

1. Show that λ is a probability on Ω^2.
2. Let $A, B \subseteq \Omega$. Show that $\lambda(A \times B) = \mu(A) \cdot \mu(B)$.

Exercise 13 — Let us treat the set $\Omega = \{1, \ldots, 1001\}$ as a discrete probability space with uniform probability.

1. Calculate $\Pr[\{k \in \Omega : 2|k \vee 3|k \vee 5|k\}]$.
2. Calculate $\Pr[\{k \in \Omega : 4|k \vee 10|k\}]$.
3. Calculate $\Pr[\{k \in \Omega : 4|k \wedge 6|k\}]$.
4. Check the correctness of your answers using Scala. Hint: use method val L = List.range(1,1002) to produce the list of all integers from 1 to 1001; then use methods L. filter(..); then use method length.

2 Probabilities

Exercise 14 — Prove that for any sequence of events $A_1, \ldots, A_n$ we have

$$\Pr[\bigcup_{i=1}^{n} A_i] = \sum_{k=1}^{n} (-1)^{k+1} \sum_{T \in [n]^k} \Pr[\bigcap_{i \in T} A_i] .$$

Remark: $[n]^k$ denotes the family of subsets of $\{1, \ldots, n\}$ with k elements (of cardinality k).

Exercise 15 — Let $\Omega = \{(x, y) \in \mathbb{R}^2 : x^2 + y^2 \leq 1\}$, $S = \mathrm{Bor}(()\Omega)$ and $\Pr(A) = \lambda(A)/\pi$, where λ is the Lebesque measure on the plane $\mathbb{R}^2$. Calculate probabilities of the followin events.

1. $A = \{(x, y) \in \Omega : x > 0\}$
2. $B = \{(x, y) \in \Omega : \frac{1}{3} < \sqrt{x^2 + y^2} \leq \frac{1}{2}\}$
3. $C = \{(x, x) \in \mathbb{R}^2 : -\frac{\sqrt{2}}{2} \leq x \leq \frac{\sqrt{2}}{2}\}$

Exercise 16 — Let $\Omega = [0, 1]$, $\mathcal{S} = \mathrm{Bor}([0, 1])$ and

$$\Pr(A) = \int_0^1 2(1 - x) \cdot \mathbf{1}_A(x)\, dx .$$

1. Check that Pr is a probability
2. Calculate $\Pr([0, \frac{1}{2}])$, $\Pr((0, \frac{1}{2}))$, $\Pr([\frac{1}{2}, 1])$.

Exercise 17 — Let $\mathcal{P} = (\Omega, \mathcal{S}, \Pr)$ be a probability space. Let $(A_n)_{n\in\mathbb{N}}$ be a sequence of elements from σ-field $\mathcal{S}$. Show that

1. $(\bigcup_{i\in\mathbb{N}} A_n)^c = \bigcap_{n\in\mathbb{N}} A_n^c$ and $(\bigcap_{i\in\mathbb{N}} A_n)^c = \bigcup_{n\in\mathbb{N}} A_n^c$.
2. $\bigcap_{i\in\mathbb{N}} A_n \in \mathcal{S}$.
3. if $A_0 \subseteq A_1 \subseteq A_2 \ldots$ then $\Pr[\bigcup_{i\in\mathbb{N}} A_n] = \lim_{n\to\infty} \Pr[A_n]$.
4. if $A_0 \supseteq A_1 \supseteq A_2 \ldots$ then $\Pr[\bigcap_{i\in\mathbb{N}} A_n] = \lim_{n\to\infty} \Pr[A_n]$.

* **Exercise 18 —** Let $\Omega = [0,1]$, $\mathcal{S} = \mathrm{Bor}([0,1])$ and let $\Pr(A)$ be the Lebesque measure of A. Let $B = \mathbb{Q}\cap[0,1]$. Show that $\Pr(B) = 0$ and $\Pr(B^c) = 1$.

Exercise 19 — Calculate the probability that among the randomly selected 13 out of 52 cards there are exactly two kings and one ace. What is the probability of having exactly one ace if you know you have exactly two kings?

Exercise 20 — What is the probability of winning a six in a „totolotek"?

3 Independent events

Exercise 21 — Suppose that events A, B, C are independent.

1. Show that events A^c, B, C are independent, too.
2. Conclude from this that events A^c, B^c, C^c independent.
3. Make a generalization of the above observations into any family of sets.

Exercise 22 — Show that an event A is independent to itself if and only if $\Pr[A] = 1$ or $\Pr[A] = 0$.

Exercise 23 — Suppose that p is a prime number. We treat $\Omega = \{1, \ldots, p\}$ as a combinatorial probability space. Show that if $A, B \subseteq \Omega$ are independent, then one of its events is equal to $\emptyset$ or Ω.

Exercise 24 — (Kolmogorov Cube) We are considering a quadrilateral dice with walls painted in the colours R, G, B and $\{R, G, B\}$. We throw this cube. Let R_X mean that the cube will fall on a wall containing the color X. Show that the events $\{R_A, R_B, R_G\}$ are pairwise independent but not independent.

Exercise 25 — We throw a coin n times. Let A_{ij} denotes the evnet „i-th and j-th throwing are the same". Show that the events $\{Aij : 1 \leq i < j \leq n\}$ are pairwise independent, but are not independent.

Exercise 26 — Show that if A and B are independent then $\Pr[A|B] = \Pr[A]$.

Exercise 27 — During the lecture, we discussed the on-line method of generating a random element from a uniform distribution.

1. Implement this method in Scala language,
2. Generate 1000 independent pointers and check the quality of the generated sample in a reasonable way.

Exercise 28 — Let $H_n = \sum_{k=1}^{n} \frac{1}{k}$. Show that

$$\ln(n+1) < H_n < 1 + \ln(n) \ .$$

Exercise 29 — Let $\Omega_{n,k}$ denotes the family of all sequences of length k from the set $\{1, \ldots, n\}$. Let $B_{n,k}$ be the set of injective sequences from $\Omega_{n,k}$. Let us treat $\Omega_{n,k}$ as a discrete probability space.

1. Derive a formula for $\Pr[B_{n,k}]$
2. Draw a plot of sequence $\{(k, B_{365,k}) : k = 1, \ldots, 365\}$
3. Find minimal k such that $B_{365,k} < 0.5$.

Exercise 30 — Population is divided into Sick (S) and healthy (H). We consider a medical test for the detection of sick people. Let T be the subpopulation for which the test gives a positive result. Suppose that $\Pr[T|S] = p$, $\Pr[T|H] = 1 - p$ and $\Pr[S] = \alpha$. Assume also that $p, \alpha \in (0,1)$.

1. Derive the formula $B(p, \alpha)$ for $\Pr[S|T]$.

2. Show that for each fixed p we have $\lim_{\alpha\to 0} B(p,\alpha) = 0$.
3. Calculate the limit $\lim_{\alpha\to 0} \frac{B(p,\alpha)}{\alpha}$.
4. Suppose that $p \approx 1$. Give a reasonable interpretation of the previous point.
5. Give interpretation of previous points.

4 Random variables

Exercise 31 — Let us consider the combinatorial probability space $\Omega = \{0,\ldots,n\}$. Niech $X(i) = i$.

1. Calculate the expected value $\mathrm{E}[X]$ of random variable X.
2. Draw the diagram of cumulative distribution function F_X of X.

Exercise 32 — Let us consider the combinatorial probability space $\Omega = \{0,\ldots,n-1\}^2$. Let $X((i,j)) = i+j$. Calculate $\mathrm{E}[X]$. Hint: Represent X as a sum of two simpler random variables.

Exercise 33 — Let F be the cumulative distribution function of X. Find cumulative distribution functions of random variables $Y = -X$, $Z = |X|$ and $U = X^2$.

Exercise 34 — Let A be an event in a probability space Ω. Let

$$\mathbf{1}_A(\omega) = \begin{cases} 1 & : \omega \in A \\ 0 & : \omega \in \Omega \setminus A \end{cases}$$

1. Show that $\mathbf{1}_A$ is a random variable
2. Calculate $\mathrm{E}[\mathbf{1}_A]$.
3. When the functions $\mathbf{1}_A$ and $\mathbf{1}_B$ are independent?
4. Do a generalization of previous result.

Exercise 35 — Let (A_n) be a partition of probability space into events. Let (x_n) be a sequence of real numbers. We put $X = \sum_n x_n \cdot \mathbf{1}_{A_n}$.

1. Show that X is a random variable.
2. Calculate $\mathrm{E}[X]$.

Exercise 36 — Check that the function $F(x) = e^{-e^{-x}}$ is a distribution of some random variable X.

1. Calculate its density function.
2. Calculate the density function of random variable $Y = X^2$.

Exercise 37 — Let $n,k \in \mathbb{N}$. Let us consider the combinatorial probability space $\Omega_{k,n}$ of all function $f : \{1,\ldots,k\} \to \{1,\ldots,n\}$.

1. Give the interpretation of this space in terms balle and urns.
2. Let $i \in \{1,\ldots,n\}$ and $E_i = \{f \in \Omega_{k,n} : i \notin rng(f)\}$. Show that $\Pr[E_i] = \left(1-\frac{1}{n}\right)^k$.
3. Let $R(f) = |rng(f)|$. Calculate $\mathrm{E}[R]$.
4. Check the last point for the case $n = k$ and calculate asymptotic of this result when $n \to \infty$.

Exercise 38 — Let us assume that X has continuous monotonic cumulative distribution F.

1. Find the distribution of random variable $Y = F(X)$.
2. Give an interpretation of previous result.

Exercise 39 — We know that if X and Y are independent random variables then $\mathrm{E}[X\cdot Y] = \mathrm{E}[X]\cdot\mathrm{E}[Y]$. Show that the assumption of independence in this result is necessary.

5 Variance

Exercise 40 — Suppose that $X \sim \mathrm{Ber}(p)$. Calculate $\mathrm{E}[X]$ and $\mathrm{var}[X]$.

Exercise 41 — Let X be a random variable and let $a, b \in \mathbb{R}$. Show that $\mathrm{var}[aX + b] = a^2 \mathrm{var}[X]$.

Exercise 42 — Calculate variance of a random variable with geometric distribution.

Exercise 43 — Let X be a random variable with exponential distribution with parameter λ (i.e

$$\Pr[X > x] = e^{-\lambda x}.$$

1. Chack that $\mathrm{E}[X] = \frac{1}{\lambda}$
2. Check that $\mathrm{var}[X] = \frac{1}{\lambda^2}$.

Exercise 44 — Let X be a random variable with exponential distribution with parameter λ. Let $x, y > 0$.

1. Show that $\Pr[X > x + y | X > x] = P[X > y]$
2. Give some reasonable interpretation of this fact

Exercise 45 — Let X be a random variable with uniform distribution on the interval $[a, b]$, i.e. that

$$F_X(t) = \begin{cases} 0 & : x < a \\ \frac{x-a}{b-a} & : x \in [a, b] \\ 1 & x > b \end{cases}$$

1. Calculate the density function of X.
2. Calculate $\mathrm{E}[X]$, $\mathrm{E}\left[X^2\right]$, $\mathrm{E}\left[X^3\right]$ and $\mathrm{var}[X]$.

Exercise 46 — Calculate $\mathrm{E}\left[(X - EX)^3\right]$ in terms $\mathrm{E}[X]$, $\mathrm{E}\left[X^2\right]$ and $\mathrm{E}\left[X^3\right]$.

Exercise 47 — The point starts from the beginning of the real line and shifts one unit to the left with a probability of 0.5 and one unit to the right with probability 0.5. Assuming that individual shifts is independent. Let D_n be the location of this point after n shifts.

1. Determine the distribution of D_n
2. Calculate $\mathrm{E}[D_n]$
3. Calculate $\mathrm{var}[D_n]$

Exercise 48 — Let B_n be the unit ball in n-dimensional space. Let X be a randomly choosen point from B_n (we consider uniform distribution). Let Y be the distance of X from the center of ball B_n. We know that $\mathrm{E}[Y] = \frac{n}{n+1}$. Calculate $\mathrm{var}[Y]$.

Exercise 49 — Let C_n be the value of Morris probabilistic counter after n ticks. We know that $\mathrm{E}\left[2^{C_n}\right] = n + 1$. Calculate $\mathrm{var}\left[2^{C_n}\right]$.

Exercise 50 — We say that a random varible X has Poisson distribution with parameter $\lambda > 0$ if it takes values in natural numbers and

$$\Pr[X = k] = \frac{e^{-\lambda} \lambda^k}{k!}$$

1. Check that this definition is correct (i.e. that $\sum_{k=0}^{\infty} \frac{e^{-\lambda} \lambda^k}{k!} = 1$).
2. Calculate $\mathrm{E}[X]$ oraz $\mathrm{var}[X]$.

Exercise 51 — Let R be the IQ of A randomly met person on the street. Suppose, that we know that $\mathrm{E}[R] = 100$ and $\mathrm{var}[R] = 100$.

1. Apply Markov inequality for upper estimation of $\Pr[X \geq 200]$.
2. Apply Thebyshev inequality for upper estimation of $\Pr[X \geq 200]$.

6 Probability generating functions

Exercise 52 — Calculate the probability generating function of the random variable with Poisson distribution and use this function for calculation expected value and variance of such random variables.

Exercise 53 — We throw a fake coin in which the probability of throwing the heads is p until we get five heads.

1. Determine the probability generating function for this experiment.
2. Calculate the expected value and variance of such random variables
3. Try to make a generalization of this exercise.

Exercise 54 — Let $N, X_1, X_2, \ldots$ be independent with values in $\mathbb{N}$. Suppose that $X_i \sim X$ for all $i \in \mathbb{N}$. Let

$$Z = \sum_{k=1}^{N} X_k$$

Calculate $\mathrm{var}\left[Z\right]$.

Exercise 55 — We consider the branching process generated by random variable $X : \Omega \to \mathbb{N}$. Let X_n denotes that size of the nth population generation. We showed during lecture that

$$\phi_{X_{n+1}} = \phi_{X_n} \circ \phi_X.$$

1. Show that $\mathrm{E}\left[X_n\right] = \left(\mathrm{E}\left[X\right]\right)^n$
2. Show that

$$\lim_{n\to\infty} \mathrm{E}\left[X_n\right] = \begin{cases} \infty & : \mathrm{E}\left[X\right] > 1 \\ 1 & : \mathrm{E}\left[X\right] = 1 \\ 0 & : \mathrm{E}\left[X\right] < 1 \end{cases}$$

3. Let $T_n = X_0 + X_1 + \ldots + X_n$ (total population size). Calculate $\mathrm{E}\left[T_n\right]$ and $\lim_{n\to\infty} \mathrm{E}\left[T_n\right]$. the branching process is generated with random variable

Exercise 56 — Consider the branching process generated by random variable with distribution $\mathrm{Bin}(3, p)$.

1. For what values of p the probability of extinction is 1?
2. Suppose that p is such that the extinction probability is less than 1. Derive a formula for the extinction probability. Draw a diagram of the generating function.
3. Suppose that $p = \frac{1}{2}$. Calculate the extiction probability.

Exercise 57 — Solve a similar exercise to the previous one for the distribution $\mathrm{Bin}(2, p)$. Explain obtained results.

7 Chernoff bounds

Exercise 58 — Calculate the moment generation functions for geometric distribution.

Exercise 59 — Compare Chernoff bounds and Thebyshev bounds for independent family of random variables $X_1, \ldots, X_n \sim \mathrm{Ber}(\frac{1}{2})$. Find n such that Chernoff gives a better result than the second one.

Exercise 60 — We proved during lectures Chernoff bounds for sequences of independent random variables $X_1, \ldots, X_n \sim \mathrm{Ber}(\frac{1}{2})$. Suppose now that $X_1, \ldots, X_n$ are independent and $X_i \sim \mathrm{Ber}(p_i)$. Let $Y_n = X_1 + \ldots + X_n$ and let $\mu = \mathrm{E}\left[Y\right]$.

1. Show that $\mu = p_1 + \ldots + p_n$.
2. Derive the Chernoff bounds for the random variable Y_n. Hint: Follow the lines of proof from the lecture. At some moment you should observe that you are done - the rest of calculation will depend only on μ and δ..

TBC

Good Luck,
Jacek Cichoń