# TD-Probability

Marco Biroli

September 28, 2020

# Chapter 1

# TD 1

## 1.1 A strategic choice.

Let $X \in \{0,1\}^3$ (resp. $Y$) be the random variable corresponding to the results of the matches using the first strategy (resp. the second strategy). Then we have that (let $D = \{(1,1,1),(1,1,0),(0,1,1)\}$):

$$P(X \in D) = a^2 b + ab(1-a) + (1-a)ba = ab(2-a)$$

Similarly:

$$P(Y \in D) = b^2 a + ba(1-b) + (1-b)ab = ba(2-b)$$

Then since $a > b$ we have that $P(X \in D) < P(Y \in D)$, hence the winning strategy is BAB.

## 1.2 Derangements

### 1.2.1

Let $E$ be a finite set and $A, B \subseteq E$. We denote by $1_A$ the indicator function of $A$ and $\bar{A}$ the complement of $A$. Then we have that:

$$1_{\bar{A}} = 1 - 1_A \quad \text{and} \quad 1_{A\cap B} = 1_A \cdot 1_B \quad \text{and} \quad 1_{A\cup B} = 1_A + 1_B - 1_{A\cap B}$$

### 1.2.2

We will prove this by induction on $n$. The base case $n = 1$ as well as $n = 2$ are trivially satisfied. Now assume that this is satisfied for $n$ then we have that (using the induction hypothesis for $n = 2$):

$$\operatorname{card}\left(\bigcup_{i=1}^{n} A_i \bigcup A_{n+1}\right) = \operatorname{card}\left(\bigcup_{i=1}^{n} A_i\right) + \operatorname{card}(A_{n+1}) - \operatorname{card}\left(\left(\bigcup_{i=1}^{n} A_i\right) \bigcap A_{n+1}\right)$$

Now we develop the last term into:

$$\left(\bigcup_{i=1}^{n} A_i\right) \bigcap A_{n+1} = \bigcup_{i=1}^{n} \left(A_i \bigcap A_{n+1}\right)$$

Now applying the induction hypothesis gives the desired result.

### 1.2.3

Let $A_i$ be the set of permutations that fixes point $i$. Then from the inclusion-exclusion principle we have:

$$D_n = n! - \operatorname{card}\left(\bigcup_{i=1}^{n} A_i\right) = n! - \sum_{k=1}^{n} (-1)^{k-1} \binom{n}{k} (n-k)! = n! \sum_{k=2}^{n} \frac{(-1)^k}{k!}$$

### 1.2.4

The probability that no one gets their jacket corresponds to the probability of having a derangement in other words:

$$p_n = \frac{D_n}{n!} = \sum_{k=2}^{n} \frac{(-1)^k}{k!} \overset{n\to\infty}{\longrightarrow} \frac{1}{e}$$

### 1.2.5

We have that:

$$D_{n,l} = \binom{n}{l} D_{n-l} = \binom{n}{l}(n-l)! \sum_{k=2}^{n-l} \frac{(-1)^k}{k!} = \frac{n!}{l!} \sum_{k=2}^{n-l} \frac{(-1)^k}{k!}$$

Hence the probability that eaxctly $l$ people leave with their jackets is:

$$p_l = \frac{D_{n,l}}{n!} = \frac{1}{l!} \sum_{k=2}^{n-l} \frac{(-1)^k}{k!}$$

### 1.2.6

The probability that a given person gets back their jacket is $p_s = \frac{1}{n}$. The probability that at least one person gets back their jacket is:

$$p_a = 1 - p_n = 1 - \sum_{k=2}^{n} \frac{(-1)^k}{k!}$$

Notice that $p_s < p_a$.

## 1.3 Balls in bins

### 1.3.1

(a) If all the balls are distinguishable then we have $\Omega = [\![1, n]\!]^r$ is the set of tuples where each element corresponds to where the $i$-th ball has been sent to. Then $\mathcal{F} = \mathcal{P}(\Omega)$ and since each event is sampled uniformly at random we have that:

$$\forall \omega \in \Omega, P(\omega) = \frac{1}{|\Omega|} = \frac{1}{n^r}$$

Then the probability of $(r_1, \cdots, r_n)$ is given by:

$$P[(r_1, \cdots, r_n)] = P[\{\omega \in \Omega : \forall i \in [\![1, n]\!] \, \#\{b \in \Omega : b = i\} = r_i\}] = \frac{1}{n^r} \binom{r}{r_1, r_2, \ldots, r_n}$$

(b) Now we have that $\Omega = \{(r_i \in \mathbb{N} : i \in [\![1, n]\!]) : \sum_{i=1}^n r_i = r\}$. Again we have that $\mathcal{F} = \mathcal{P}(\Omega)$. Then we have that:

$$P[(r_1, \ldots, r_n)] = \frac{1}{|\Omega|} = \frac{1}{\binom{r+n-1}{n-1}}$$

(c) Now we have that $\Omega = \{s \in \{0, 1\}^n : \sum_{i=1}^n s_i = r\}$ corresponding to the tuple indicating if each state is occupied or not. Once again $\mathcal{F} = \mathcal{P}(\Omega)$. Now the probability is given by:

$$P[(r_1, \ldots, r_n)] = \frac{1}{|\Omega|} = \frac{1}{\binom{n}{r}}$$

### 1.3.2

The probability that at least two have the same birthday is the 1 minus the probability that none of them share a birthday. The probability that none of them share a birthday is given by $\frac{r!}{n^r}\binom{n}{r}$. Hence the probability that at least two people share a birthday is given by: $1 - \frac{r!}{n^r}\binom{n}{r}$.

### 1.3.3

Days are bins, accidents are distinguishable balls hence the probability is given by:

$$\frac{\binom{r}{n} n^{r-n}}{n^r} = n^{-n} \binom{r}{n}$$

### 1.3.4

# Chapter 2

# TD2

## 2.1 Symmetric Random Walk

Consider a balanced coin drawn $n$ times. Denote $X_1, \cdots, X_n$ the results and $S_k$ the partial sums.

### 2.1.1

The law of $S_k$ is given by:

$$p_{n,r} = P(S_n = r) = \frac{1}{2^n}\binom{n}{\frac{n+r}{2}}$$

### 2.1.2

The number of paths from $(0,0)$ to $(2n+2,0)$ never zero are equal to the number of paths from $(1,1)$ to $(2n+1,1)$ which always stay above or equal to the line $y=1$. Hence rescaling the $y$-axis by a factor 1 we get a bijection in between the strictly positive walks from $(0,0)$ to $(2n+2,0)$ with the positive or zero walks from $(0,0)$ to $(2n,0)$. Hence for symmetric random walks the number of random walks going from $(0,0)$ to $(2n+2,0)$ never touching the axis is twice as much as the number of walks from $(0,0)$ to $(2n,0)$ being always positive or 0. Furhtemore there are 4 times more walks going from 0 to $2n+2$ which therefore gives the desired result.

### 2.1.3

We now that the end of the random walk is going to be given by $a-b$. Now the number of possible only positive walks is given by the number of walks from $(1,1)$ to $(a+b, a-b)$ minus the number of walks from $(1,-1)$ to $(a+b, a-b)$ by the reflexion principle. Hence we get that:

$$p = p_{a+b-1,a-b-1} - p_{a+b-1,a-b+1} = \frac{1}{2^n}\frac{a-b}{a+b}\binom{a+b}{a}$$

### 2.1.4

#### a)

Up to a re-scaling of the $y$-axis we have the equivalent problem of computing the number of paths that go from $(0,-r)$ to $(n, k-r)$ and which touch the $x$-axis at least once. Now notice that from the reflexion principle this is equal to the number of paths from $(0,r)$ to $(n,k-r)$ and up to a second shifting this is equal to the number of paths from $(0,0)$ to $(n,k-2r)$. Hence the desired probability is given by $p_{n,k-2r} = p_{n,2r-k}$.

**b)**

Then for any $r$ we have that:

$$\begin{aligned}
P(\max\{S_1, \cdots, S_n\} = r) &= \sum_{k=-\infty}^{+\infty} P(S_n = k, \max\{S_1, \cdots, S_n\} = r) \\
&= \sum_{k=-\infty}^{+\infty} (P(S_n = k, \max\{S_1, \cdots, S_n\} \geq r) - P(S_n = k, \max\{S_1, \cdots, S_n\} > r)) \\
&= \sum_{k=-\infty}^{+\infty} (P(S_n = k, \max\{S_1, \cdots, S_n\} \geq r) - P(S_n = k, \max\{S_1, \cdots, S_n\} \geq r+1)) \\
&= \sum_{k=-\infty}^{+\infty} (p_{n,k-2r} - p_{n,k-2r-2}) = p_{n,r} + p_{n,r+1}
\end{aligned}$$

**c)**

This can be re-written as:

$$P(S_n = 0, S_1 < 0, \cdots, S_{n-1} < 0, S_0 = -r) = P(S_n = -r, S_1 < 0, \cdots, S_{n-1} < 0, S_0 = 0)$$

Now again notice that this corresponds to a symmetric re-writing of the problem 3. Hence we get immediately that:

$$\hat{p}_{n,r} = P(S_n = r, S_1 < r, \cdots, S_{n-1} < r) = \frac{1}{2^n}\frac{r}{n}\binom{n}{\frac{n+r}{2}} = \frac{r}{n}p_{n,r}$$

**d)**

## 2.2 Geometric and negative-binomial laws.

### 2.2.1

We want to compute:

$$P(T_1 = t+1)$$

For such a thing to be the case we need to have thrown tails successively $t$ times and heads the last time. Hence the probability is given by:

$$P(T_1 - 1 = t) = P(T_1 = t+1) = (1-p)^t p = \mathcal{G}(p)$$

The expectancy of $\mathcal{G}(p)$ is given by:

$$\mathbb{E}[\mathcal{G}(p)] = \sum_{t=0}^{+\infty}(1-p)^t pt = p \cdot \frac{1-p}{p^2} = \frac{1-p}{p}$$

The variance of $\mathcal{G}(p)$ is given by:

$$\mathrm{Var}[\mathcal{G}(p)]^2 = \frac{p^2 - 3p + 2}{p^2} - \frac{1 - 2p + p^2}{p^2} = \frac{1-p}{p^2}$$

### 2.2.2

A geometric law corresponds to something not happening $t$ times and then happening at the $t+1$ time. Then the infimum of two geometric laws corresponds to two things not happening $t$ times and one of them happening at the $t+1$ time. The probability of which is given as follows:

$$P[(\inf(S_1, S_2) = s] = (1-p)^{2(s-1)}(p^2 + 2p(1-p)) = (1-p)^{2(s-1)}(2p - p^2) = (1-p)^{2(s-1)}(1 - (1-p)^2)$$

Hence the infimum is a geometric variable with law $\mathcal{G}(1 - (1-p)^2)$

### 2.2.3

We want to compute $P(T_m - m = k)$. The number of possible outcomes for which $T_m - m = k$ is given by $\binom{k+m-1}{m-1}p^m(1-p)^k = Neg(m,p)$. Now using the formula we get that:

$$\sum_{k\geq 0} Neg(m,p)[k] = p^m(1-(1-p))^{-m} = 1$$

### 2.2.4

Notice that we can re-write $T_m - m$ as the number of steps before the first $H$ plus the number of steps between the first and second head etc. now the number of steps in between two succesive heads is given by $\mathcal{G}(p)$ then the result follows.

### 2.2.5

Then we have that:

$$\mathbb{E}[T_m - m] = m\mathbb{E}[\mathcal{G}(p)] = m\frac{1-p}{p}$$

Hence:

$$\mathbb{E}[T_m] = \frac{m}{p}$$

Similarly we get:

$$Var(T_m) = m\frac{1-p}{p^2}$$

## 2.3 Thinning and Poisson random variables

We have that:

$$P[Y = k|X = n] = \binom{n}{k}p^k(1-p)^{n-k}$$

Then from the law of total probability we have that:

$$P[Y = k] = \sum_{n=k}^{+\infty}\binom{n}{k}p^k(1-p)^{n-k}\frac{\lambda^n e^{-\lambda}}{n!} = \frac{p^k e^{-\lambda}\lambda^k}{k!}\sum_{n=0}^{+\infty}(1-p)^n\frac{\lambda^n}{(n)!} = \frac{(\lambda p)^k e^{-\lambda p}}{k!}$$

## 2.4 Conditional Probabilities

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

   (b) From the previous question we have:

   $$P(B|A) = \frac{P(A)}{P(B)}P(A|B) = \frac{P(A|B)}{P(B)}\sum_{i\in\mathbb{N}}P(A|C_i)P(C_i)$$

   (c) Let $X$ be the r.v. corresponding to the number of children of law $p_i$ and $A$ the event "has no daughter". Then:

   $$P(X = 1|A) = \frac{P(A|X=1)}{P(X=1)}\sum_{i\in\mathbb{N}}P(A|X=i)P(X=i) = \frac{1}{2p_1}\sum_{i\in\mathbb{N}}\frac{p_i}{2^i}$$

2. The law of $(X,Y)$ is given by:

   $$P((X,Y) = (x,y)) = \frac{1}{6}\cdot\frac{\delta_{y\leqslant x}}{x} \quad \text{for} \quad (x,y)\in[\![1,6]\!]^2$$

   Then:

   $$P(X = x) = \sum_{y=1}^{6}P((X,Y) = (x,y)) = \sum_{y=1}^{x}\frac{1}{6x} = \frac{1}{6}$$

   Similarly:

   $$P(Y = y) = \sum_{x=1}^{6}P((X,Y) = (x,y)) = \frac{1}{6}\sum_{x=y}^{6}\frac{1}{x}$$

## 2.5 Change of variables.

1. Let $U$ be a r.v. with uniform law over $[-\frac{\pi}{2}, \frac{\pi}{2}]$. Then:

$$P(\tan(U) \leq x) = P(U \leq \tan^{-1}(x)) = \frac{\arctan(x) - \frac{\pi}{2}}{\pi}$$

Then the pdf of $\tan U$ is given by:

$$f(x) = \frac{1}{\pi(1+x^2)}$$

Then the value of $\mathbb{E}[|\tan U|]$ is given by:

$$\mathbb{E}[|\tan U|] = \int_{-\infty}^{+\infty} \frac{|x|}{\pi(1+x^2)} \mathrm{d}x$$

Which diverges.

2. We have that $X = \cos\theta$ and $Y = \sin\theta$ and $\theta$ is a r.v. with uniform law on $[0, 2\pi[$. Then:

$$f_X(x) = f_{\cos\theta}(x) = \frac{1}{\pi}\left|\frac{\mathrm{d}}{\mathrm{d}x}\arccos x\right| = \frac{1}{\pi\sqrt{1-x^2}}$$

By symmetry it is immediate that $P(X = x) = P(Y = x)$. Then $z = X + Y = \cos\theta + \sin\theta = \sqrt{2}\cos(\theta - \frac{\pi}{4})$. Then:

$$f_Z(z) = f_x\left(\frac{z}{\sqrt{2}}\right)\frac{1}{\sqrt{2}} = \frac{1}{\pi}\frac{1}{\sqrt{2-z^2}}$$

## 2.6 Random Variables

1. Let $X(\omega) = a1_A(\omega) + b1_B(\omega)$ and $C \in \mathcal{B}(\mathbb{R})$. Then:

$$\sigma(X) = \langle \emptyset, A, B, A\cup B, A\cap B\rangle \text{ and } P(X = x) = \begin{cases} P(A) & \text{if } x = a \\ P(B) & \text{if } x = b \\ P(A\cap B) & \text{if } x = a+b \end{cases}$$

2. Then:

$$P_X = \frac{1}{2}\delta_1 + \frac{1}{2}Unif([0,1]) \text{ and } \sigma(X) = \langle \emptyset, \mathcal{B}([0,1/2]), [1/2,1]\rangle$$

3. Then:

$$P(X \leq x) = \int_{-\sqrt{x}}^{\sqrt{x}} \frac{\mathrm{d}x}{2} = \sqrt{x} \text{ and } \sigma(X) = \{A \in \mathcal{B}([-1,1]) : A = -A\}$$

# Chapter 3

# TD3

## 3.1 Gamma Law

1. We have that:

$$f_X(x) = f_{Z/\lambda}(x) = h_{a,1}(\lambda x)\lambda = 1_{x\geq 0}\frac{1}{\Gamma(a)}1^a(\lambda x)^{a-1}\exp(-\lambda x)\lambda = h_{a,\lambda}(x)$$

2. We have that:

$$E[Z] = \int_0^{+\infty} \frac{x^{a-1}e^{-x}}{\Gamma(a)} x\mathrm{d}x = \frac{1}{\Gamma(a)}\int_0^{+\infty} x^{(a+1)-1}e^{-x}\mathrm{d}x = \frac{\Gamma(a+1)}{\Gamma(a)} = a+1$$

Which also immediately gives:

$$E[X] = E[Z/\lambda] = \frac{a}{\lambda}$$

Then the variance is given by a similar integration which yields:

$$Var[Z] = \frac{\Gamma(a+2)}{\Gamma(a)} - \frac{\Gamma(a+1)}{\Gamma(a)} = a$$

Which also gives:

$$Var[X] = Var[Z/\lambda] = \frac{a}{\lambda^2}$$

3. Using question 1 it is sufficient to show the case where $X, Y$ have laws $\mathcal{G}(a,1)$ and $\mathcal{G}(b,1)$. Then:

$$f_Z(z) = f_X \star f_Y(z) = \int f_{X,Y}(x, z-x)\mathrm{d}x = \int f_X(x)f_Y(z-x)\mathrm{d}x$$

Now replacing with the laws we get:

$$f_Z(z) = \int_0^z \frac{1}{\Gamma(a)\Gamma(b)} x^{a-1}(z-x)^{b-1}e^{-\lambda x-\lambda(z-x)}\mathrm{d}x = \frac{z^{a-1+b-1+1}e^{-\lambda z}}{\Gamma(a)\Gamma(b)}\int_0^1 u^{a-1}(1-u)^{b-1} = h_{a+b,1}(z)$$

Where the last equality follows from normalization. For a proof check the next question.

## 3.2 Beta law

1. We compute:

$$\int_0^1 x^{a-1}(1-x)^{b-1}\mathrm{d}x =$$

2. We introduce $V = X$ and make the change of variable $(X,Y) \to (Z,V)$. Which explicitly gives $(Z,V) = (XY, X)$ Then the Jacobian determinant is given by:

$$\begin{vmatrix} y & x \\ 1 & 0 \end{vmatrix} = -x = -v$$

Then we have:

$$f_{Z,V}(z,v) = f_{x,y}(v, \frac{z}{v})\frac{1}{|v|}$$

And integrating gives:

$$f_Z(z) = \int f_{X,Y}(x, \frac{z}{x})\frac{1}{|x|}\mathrm{d}x = \int f_X(x) f_Y(\frac{z}{x})\frac{1}{|x|}\mathrm{d}x$$

Then plugging in the laws gives the desired result.

## 3.3 Gaussian Law

1. We have:

$$E[X] = \frac{1}{\sqrt{2\pi}}\int_{-\infty}^{+\infty} \exp\left[-\frac{x^2}{2}\right] x\mathrm{d}x = 0$$

   Since the integrand is odd. Then the variance is given by:

$$E[X^2] = \frac{1}{\sqrt{2\pi}}\int_{-\infty}^{+\infty} \exp\left[-\frac{x^2}{2}\right] x^2\mathrm{d}x = 1$$

2. The law of $Y = m + \sigma X$ is given by:

$$f_Y(y) = f_X(\frac{y-m}{\sigma})\frac{1}{\sigma} = \frac{1}{\sqrt{2\pi}\sigma}\exp\left[-\frac{(\frac{y-m}{\sigma})^2}{2}\right] = h_{m,\sigma^2}(y)$$

   Then we have that:

$$E[Y] = m \text{ and } Var[Y] = \sigma^2$$

3. Let $Z = X/Y$ and $V = Y$ then the Jacobian determinant is given by:

$$\begin{vmatrix} \frac{1}{Y} & -\frac{X}{Y^2} \\ 0 & 1 \end{vmatrix} = \frac{1}{Y}$$

   Then we have that:

$$f_{Z,V}(z,v) = f_{X,Y}(zv, v)|y| \Rightarrow f_Z(z) = \int f_X(zv) f_Y(v)|v|\mathrm{d}v$$

   Which when replacing with the laws gives:

$$f_Z(z) = \frac{1}{\pi}\frac{1}{1+z^2}$$

   From symmetry the law of $1/Z$ is identical.

# Chapter 4

# TD 4

## 4.1 Around the notion of independence.

1. $A$ and $B$ are independent, $B$ and $C$ also and $A$ and $C$ also. However $A, B, C$ is not independent.

2. $P(A) = P(A \cap A) = P(A)P(A) = P(A)^2$ then $P(A) = 0$ or $P(A) = 1$. Similarly $P(X = a) = P(X = a|X = b) = \delta_{ab}$. So $X$ must be single valued and have that value with probability 1.

3. We have that $E[X] = E[Y] = 0$. Since $X, Y$ are independent we also have that $E[Z] = E[XY] = E[X]E[Y] = 0$. Hence we have that:

$$Cov[X, Z] = E[(X - E[X])(Z - E[Z])] = E[XZ] = E[X^2Y] = E[X^2]E[Y] = E[X^2] \cdot 0 = 0$$

$X$ and $Z$ are not independent because:

$$P(Z = 0|X = 0) = 1 \neq P(Z = 0)$$

## 4.2 Convergence in probability of random variables.

1. We have that:

$$\forall \varepsilon > 0, \lim_{n \to +\infty} P(|X_n - X| > \varepsilon) = 0$$

Now notice that:

$$P(|YX_n - YX| > \varepsilon) = P(|Y||X_n - X| > \varepsilon)$$

Since this must be true for all $\varepsilon$ it is equivalent to saying that:

$$P(|X_n - X| > \varepsilon \wedge |Y| > \varepsilon) \leq P(|X_n - X|)$$

Which by hypothesis converges to 0.

2. Fix $\varepsilon$ then:

## 4.3 Different modes of convergence of random variables.

1. $(X_n)$ converges to $\delta_0$ almost surely from construction. Then we have that:

$$E[|X_n - \delta_0|] = E[|X_n|] = 0 \cdot (1 - \frac{1}{n^2}) + n^2 \cdot \frac{1}{n^2} = 1$$

Hence it does not converge to $\delta_0$ in $L^1$.

2. (a) We have that:

$$E[X_n] = 0(1 - \frac{1}{n}) + 1 \cdot \frac{1}{n} = \frac{1}{n} \overset{n \to +\infty}{\longrightarrow} 0$$

Hence $X_n$ goes to 0 in $L^1$ and therefore also in probability.

(b) We have that:

$$P(\limsup A_n) = P(\bigcap_m \bigcup_{n \geq m} A_n) = P(\lim_{n \to +\infty} \exists m \geq n, X_m = 1)$$