Sheet 11

Please put your solution in the box of your corresponding instructor before **Friday, May 15th** at **12h00** and name your sheet “<First name>_<Last name>.pdf”.

Exercise 1. We let $p \in (0,1)$, and, for every $n \in \mathbb{N}$, X_n be a random variable following a Binomial distribution with parameters (n,p).

1. For all $\lambda \in \mathbb{R}$ and $n \in \mathbb{N}$, give a simple formula for $\mathbb{E}[e^{\lambda X_n}]$. Prove that, for all $x \geq 0$,
$$\mathbb{P}[X_n \geq nx] \leq \exp\Big[n\Big(\log(1+p(e^\lambda - 1)) - \lambda x\Big)\Big].$$
2. Suppose $1 > x > p$. Optimize the above inequality over λ (taking the necessary precautions) and find a function I such that, if $x > p$,
$$\mathbb{P}[X_n \geq nx] \leq \exp(nI(x)).$$
Study the function I; in particular, discuss the similarities and differences with the statement of Cramér's theorem.
3. If $x < p$, how does one study $\mathbb{P}[X_n \leq nx]$?

Solution of exercise 1.

1. We use the well-known fact that a Binomial distribution with parameters (n,p) can be realised as the sum of n independent Bernoulli's of parameter p. Let then $(Y_k)_{k\in\mathbb{N}}$ be a sequence of independent Bernoulli's of parameter p; then since
$$X_n \sim \sum_{k=1}^{n} Y_k$$
one has
$$\mathbb{E}[e^{\lambda X_n}] = \mathbb{E}[e^{\lambda \sum_{k=1}^n Y_k}] = \prod_{k=1}^{n} \mathbb{E}[e^{\lambda Y_k}] = \Big(1 + p(e^\lambda - 1)\Big)^n,$$
where we have used the independence of the Y_k's to exchange product and expectation.

   Now, we use Markov's inequality:
$$\mathbb{P}[X_n \geq nx] = \mathbb{P}[e^{\lambda X_n} \leq e^{\lambda nx}] = \frac{\Big(1+p(e^\lambda - 1)\Big)^n}{e^{\lambda nx}},$$
which we can simplify as
$$\mathbb{P}[X_n \geq nx] \leq \exp\Big[n\Big(\log(1+p(e^\lambda - 1)) - \lambda x\Big)\Big].$$

2. The function $\lambda \mapsto e^{\lambda}$ is C^{∞} on $\mathbb{R}$, and the function $t \mapsto \log(t)$ is C^{∞} on $(0,+\infty)$. Here, one has $1+p(e^{\lambda}-1) > 1-p \leq 0$ for all $\lambda \in \mathbb{R}$, so that the function $f : \lambda \mapsto \log(1+p(e^{\lambda}-1)) - \lambda x$ belongs to $C^{\infty}(\mathbb{R},\mathbb{R})$.

   By the chain rule,
$$f' : \lambda \mapsto \frac{pe^{\lambda}}{1+p(e^{\lambda}-1)} - x$$
   vanishes at exactly one point:
$$\lambda_0 = \log\left(\frac{x(1-p)}{p(1-x)}\right).$$
   To prove that this point is a minimum, we study the behaviour of f near $\pm\infty$; near $-\infty$, one has $\log(1+p(e^{\lambda}-1)) \to \log(1-p)$ so that $f \to +\infty$. Near $+\infty$, one has $\log(1+p(e^{\lambda}-1)) \sim \lambda$, so that, since $x<1$, $f \to +\infty$. In conclusion, f is minimal at λ_0, and we obtain the desired result with
$$\begin{aligned} I(x) = f(\lambda_0) = \log\left[1+p\left(\frac{x(1-p)}{p(1-x)}-1\right)\right] - \log\left(\frac{x(1-p)}{p(1-x)}\right)x \\ = (1-x)\log\left(\frac{x(1-p)}{p(1-x)}\right). \end{aligned}$$
   To see the link with Cramér's theorem, observe that since
$$\log(1+p(e^{\lambda}-1)) = \log \mathbb{E}[e^{\lambda Y_1}]$$
   is finite for every $\lambda \in \mathbb{R}$, Cramér's theorem applies. We already identified $I(x) = \inf_{\lambda\in\mathbb{R}}(\log \mathbb{E}[e^{\lambda Y_1}] - \lambda x)$ so that, by Cramér's theorem, one has in fact, for every fixed $x \in (p,1)$, as $n \to +\infty$,
$$\mathbb{P}[X_n \geq nx] \sim \exp(-nI(x)).$$
   So what we just proved, in this particular example, is one half of Cramér's theorem.

3. There is a symmetry in Bernoulli (hence, binomial) random variables: $Y_i^* = 1-Y_i$ is a Bernoulli with parameter $1-p$, so that $X_n^* = n - X_n$ is a Binomial with parameters $(n, 1-p)$. In particular,
$$\mathbb{P}[X_n \leq nx] = \mathbb{P}[X_n^* \geq n(1-x)]$$
   can be estimated as previously, replacing x with $1-x$ and p with $1-p$, since $1-x > 1-p$.

□

Exercise 2. Let $(\mu_n)_{n\in\mathbb{N}}$ be a sequence of probability measures on $\mathbb{N}$: for every n, μ_n is a sequence $(\mu_{n,k})_{k\in\mathbb{N}}$ of elements of $[0,1]$, such that $\sum_{k=1}^{+\infty} \mu_{n,k} = 1$ for every n.

1. Suppose that $(\mu_n)_{n\in\mathbb{N}}$ converges[1] to a probability measure $\overline{\mu}$. Show that
   a) For every $k \in \mathbb{N}$, as $n \to +\infty$, one has $\mu_{n,k} \to \overline{\mu}_k$.
   b) For every $\epsilon > 0$, there exists N such that, for every $n \in \mathbb{N}$,
   $$\mu_n(\{1,...,N\}) \geq 1-\epsilon.$$
   Hint: first prove b) for n large enough (depending on ϵ but not on N), then increase N to get all $n \in \mathbb{N}$.
2. Suppose reciprocally that there exists a sequence $\overline{\mu}_k$ such that conditions a) and b) above are satisfied. Prove that $\overline{\mu}$ is a probability measure and that $(\mu_n)_{n\in\mathbb{N}}$ converges to μ.
3. Suppose that condition b) above is satisfied. Show that there exists a probability sequence $\overline{\mu}$ such that you can extract a subsequence of $(\mu_n)_{n\in\mathbb{N}}$ converging towards $\overline{\mu}$. (You may use a diagonal argument).

Solution of exercise 2.

1. Let $f : \mathbb{N} \mapsto \mathbb{R}$ defined by $f(k) = 1$ and $f(j) = 0$ if $j \neq k$. Then, by definition of convergence of probability measures, as $n \to +\infty$ one has
   $$\sum_{j=0}^{+\infty} \mu_{n,j} f(j) \to \sum_{j=0}^{+\infty} \overline{\mu}_j f(j).$$
   By design, the left-hand side is equal to $\mu_{n,k}$ and the right-hand side to $\overline{\mu}_k$. This proves a).

   To prove b), we start from the definition of convergence of probability measures, applied to $f = 1$ if $k \leq N$ and 0 if $k > N$:
   $$\forall N, \forall \epsilon, \exists n_0(N,\epsilon), \forall n \geq n_0(N,\epsilon), \left|\mu_n(\{1,\dots,N\}) - \overline{\mu}(\{1,\dots,N\})\right| \leq \frac{\epsilon}{2}.$$
   By σ-additivity of $\overline{\mu}$, one has $\lim_{N\to+\infty} \overline{\mu}(\{1,\dots,N\}) = 1$, so that there exists $N_0(\epsilon)$ such that
   $$\overline{\mu}(\{1,\dots,N_0(\epsilon)\}) \geq 1 - \frac{\epsilon}{2}.$$
   Hence, for $n \geq n_0(\epsilon, N_0(\epsilon))$, one has indeed
   $$\mu_n(\{1,\dots,N_0(\epsilon)\}) \geq 1-\epsilon.$$
   However, that property may fail for $n < n_0(\epsilon, N_0(\epsilon))$. We fix ϵ, then there are only a finite number of $n \in \mathbb{N}$ such that $n < n_0(\epsilon, N_0(\epsilon))$. By σ-additivity of μ_n, for each of those n, there exists $N_1(n)$ such that, for all $N \geq N_1(n)$, one has $\mu_n(\{1,\dots,N\}) \geq 1-\epsilon$. To conclude the claim, we let
   $$N = \max(N_0(\epsilon), N_1(0), N_1(1), \dots, N_1(n(\epsilon, N_0(\epsilon))) - 1).$$

[1] Here, a bounded continous function f from $\mathbb{N}$ to $\mathbb{R}$ simply means a bounded real sequence; compact support means finite number of non-zero elements.

2. By property a), since $\mu_{n,k} \in [0,1]$ for every k, n, then $\overline{\mu}_k \in [0,1]$. To prove that $\overline{\mu}$ defines a probability measure, we need to check that $\sum_k \overline{\mu}_k = 1$. We use b) (and the fact that μ_n is a probability measure) to this end: for every $\epsilon > 0$, there exists N such that, for every $n \in \mathbb{N}$,

$$1 \geq \sum_{k=1}^{N} \mu_{n,k} \geq 1 - \epsilon.$$

Applying a), we obtain that, for all $\epsilon > 0$, there exists $N \geq 0$ such that

$$1 \geq \sum_{k=1}^{N} \overline{\mu}_k \geq 1 - \epsilon.$$

Then, by the sandwich theorem, $\sum_{k=1}^{N} \overline{\mu}_k = 1$. Let now $f : \mathbb{N} \to \mathbb{R}$ be a bounded sequence. We let $M = \sup |f|$.

As a consequence of b), for all $\epsilon > 0$, there exists $N \geq 0$ such that, for all $n \in \mathbb{N}$,

$$\left| \mu_n(\{N+1, N+2, ...\}) - \overline{\mu}(\{N+1, N+2, \ldots\}) \right| \leq 2\epsilon,$$

so that

$$\left| \sum_{k=N+1}^{+\infty} f(k)\mu_{n,k} - \sum_{k=N+1}^{+\infty} f(k)\overline{\mu}_k \right| \leq 2M\epsilon.$$

Now, by a), there exists n_0 such that, for all $n \geq n_0$, for all $k \leq N$, one has $|\mu_{n,k} - \overline{\mu}_k| \leq \epsilon$, so that, for $n \geq n_0$,

$$\left| \sum_{k=0}^{N} f(k)\mu_{n,k} - \sum_{k=0}^{N} f(k)\overline{\mu}_k \right| \leq M\epsilon.$$

To conclude, for all $n \geq n_0$, one has $|\mu_n(f) - \overline{\mu}(f)| \leq 3M\epsilon$. Since M is fixed and ϵ is arbitrary, we conclude that $\mu_n(f) \to \mu(f)$.

3. We only have to construct a subsequence of $(\mu_n)_{n\in\mathbb{N}}$ such that a) holds.

   Since $(\mu_n)_{n\in\mathbb{N}}$ is a sequence of probability measures, for every k, the sequence $(\mu_{n,k})$ is bounded (and belongs to $[0,1]$). Since $(\mu_{n,0})$ is bounded, in particular, there exists $\phi_0 : \mathbb{N} \to \mathbb{N}$ increasing such that $(\mu_{\phi_0(n),0})$ converges to some $\overline{\mu}_0$.

   Now the sequence $(\mu_{\phi_0(n),1})$ is bounded so that we can again extract a subsequence: there exists $\phi_1 : \mathbb{N} \to \mathbb{N}$ increasing such that $(\mu_{\phi_1(\phi_0(n)),1})$ converges to $\overline{\mu}_1$; by construction, $\mu_{\phi_1(\phi_0(n)),0}$ is a subsequence of $\mu_{\phi_0(n),0}$ and converges to $\overline{\mu}_0$.

   Continuing like this, for every K there exists an extraction $\psi_K = \phi_K \circ \phi_{K-1} \circ \phi_{K-2} \circ \ldots \circ \phi_0$ such that condition a) is satisfied on the subsequence $(\mu_{\psi_K(n)})$ for $0 \leq k \leq K$.

To conclude, we consider the **diagonal** subsequence

$$\psi^{\mathrm{d}iag} : n \mapsto \psi_n(n)$$

(that is, we have a double sequence depending on n and K, and we take the diagonal values $K = n$). Then $(\mu_{\psi^{\mathrm{d}iag}(n)})$ satisfies a) and b), thus converges to $\overline{\mu}$.

□

Exercise 3.

1. Using exercise sheet 9, prove rapidly that, if $(X_n)_{n\in\mathbb{N}}$ and $(Y_n)_{n\in\mathbb{N}}$ are sequence of real random variables such that $(X_n)_{n\in\mathbb{N}}$ converges in distribution to some random variable X and $(Y_n)_{n\in\mathbb{N}}$ converges (in distribution/probability) to a constant c, then
$$X_n + Y_n \Rightarrow X + c \qquad\qquad X_n Y_n \Rightarrow cX.$$

2. We let $(X_i)_{i\in\mathbb{N}}$ be a sequence of independent, identically distributed Gaussian variables of expectation μ and variance σ. We consider their *empirical* variance and mean:
$$\overline{X}_n = \frac{1}{n}\sum_{i=1}^{n} X_i \qquad\qquad V_n(X) = \frac{1}{n-1}\sum_{i=1}^{n}(X_i - \overline{X}_n)^2.$$
What are the expectations of $\overline{X}_n$ and $V_n(X)$? Show that, as $n \to +\infty$, these random variables converge (in an appropriate sense) to their expectations.

3. Show that, as $n \to +\infty$,
$$\frac{X_1 + \ldots + X_n - n\mu}{\sqrt{nV_n(X)}} \Rightarrow \mathcal{N}(0,1).$$
Can one replace μ with $\overline{X}_n$ in the result above?

Solution of exercise 3.

1. We proved in sheet 9 that $(X_n, Y_n) \Rightarrow (X, c)$. In other terms, for every continuous bounded function $f : \mathbb{R}^2 \to \mathbb{R}$, as $n \to +\infty$, there holds $\mathbb{E}[f(X_n, Y_n)] \to \mathbb{E}[f(X, c)]$.

   In particular, let $g : \mathbb{R} \to \mathbb{R}$ be continuous and bounded. Then $f_1 : (x, y) \mapsto g(x + y)$ is itself continuous and bounded so that
$$\mathbb{E}[g(X_n + Y_n)] = \mathbb{E}[f_1(X_n, Y_n)] \to \mathbb{E}[f_1(X, c)] = \mathbb{E}[g(X + c)];$$
   hence $X_n + Y_n \Rightarrow X + c$. In the same fashion,
$$f_2 : (x, y) \mapsto g(xy)$$
   is continous and bounded, and we conclude that $X_n Y_n \Rightarrow cX$.

2. By linearity of the expectation, the expectation of $\overline{X}_n$ is μ; since $E[|X_i|] \leq \infty$, by the law of large numbers, $\overline{X}_n$ converges almost surely to $E[X_1] = \mu$. Similarly, the expectation of $V_n(X)$ is

$$\begin{aligned}\mathbb{E}[V_n(X)] &= \frac{1}{n-1}\sum_{i=1}^{n}\mathbb{E}\left[\left(X_i - \frac{1}{n}\sum_{j=1}^{n}X_j\right)^2\right] \\ &= \frac{1}{n-1}\sum_{i=1}^{n}\mathbb{E}\left[X_i^2 - \frac{2}{n}\sum_{j=1}^{n}X_iX_j + \frac{1}{n^2}\left(\sum_{j=1}^{n}X_j\right)^2\right] \\ &= \frac{1}{n-1}n\mathbb{E}[X_1^2] - \frac{1}{n(n-1)}\sum_{i,j=1}^{n}\mathbb{E}[X_iX_j],\end{aligned}$$

where we develop

$$\mathbb{E}\left[\left(\sum_{j=1}^{n}X_j\right)^2\right] = \mathbb{E}\left[\sum_{i,j=1}^{n}X_iX_j\right] = \sum_{i,j=1}^{n}\mathbb{E}[X_iX_j].$$

If $i \neq j$, then X_i is independent from X_j, so that

$$\begin{aligned}\mathbb{E}[V_n(X)] &= \left(\frac{n}{n-1} - \frac{1}{n-1}\right)\mathbb{E}[X_1^2] - \frac{1}{n(n-1)}\sum_{i\neq j}\mathbb{E}[X_1]^2 \\ &= \mathbb{E}[X_1^2] - \mathbb{E}[X_1]^2 = \mathrm{Var}(X_1) = \sigma^2.\end{aligned}$$

We cannot apply the law of large numbers, so, in order to prove convergence of $V_n(X)$ towards σ^2, we will prove that the variance of $V_n(X)$ tends to zero (this proves L^2 convergence, which implies convergence in probability).

The expectation of $V_n(X)^2$ is

$$\frac{1}{(n-1)^2}\sum_{i_1=1}^{n}\sum_{i_2=1}^{n}\mathbb{E}\left[X_{i_1}^2X_{i_2}^2 - \frac{4}{n}X_{i_1}^2X_{i_2}\sum_{j_1=1}^{n}X_{j_1} + \frac{2}{n^2}\sum_{j_1=1}^{n}\sum_{j_2=1}^{n}X_{i_1}X_{i_2}X_{j_1}X_{j_2}\right.$$
$$\left. + \frac{2}{n^2}\sum_{j_1=1}^{n}\sum_{j_2=1}^{n}X_{i_1}^2X_{j_1}X_{j_2} - \frac{4}{n^3}\sum_{j_1=1}^{n}\sum_{j_2=1}^{n}\sum_{j_3=1}^{n}X_{i_1}X_{j_1}X_{j_2}X_{j_3} + \frac{1}{n^4}\sum_{j_1=1}^{n}\sum_{j_2=1}^{n}\sum_{j_3=1}^{n}\sum_{j_4=1}^{n}X_{j_1}X_{j_2}X_{j_3}X_{j_4}\right].$$

There are only three different kind of terms in this expression, so that

$$\begin{aligned}\mathbb{E}[V_n(X)^2] = \frac{1}{(n-1)^2}\sum_{i_1,i_2}\mathbb{E}[X_{i_1}^2X_{i_2}^2] - \frac{2}{(n-1)^2n}\sum_{i_1,i_2,i_3}\mathbb{E}[X_{i_1}^2X_{i_2}X_{i_3}] \\ - \frac{1}{n^4}\sum_{i_1,i_2,i_3,i_4}\mathbb{E}[X_{i_1}X_{i_2}X_{i_3}X_{i-4}].\end{aligned}$$

We let $C_3 = \mathbb{E}[X_1^3]$ and $C_4 = \mathbb{E}[X_1^4]$ be the third and fourth moments of $\mathbb{N}(\mu, \sigma^2)$. Then, with $\mathbb{E}[X_1^2] = \sigma^2 + \mu^2$, one has

$$\frac{1}{(n-1)^2} \sum_{i_1, i_2} \mathbb{E}[X_{i_1}^2 X_{i_2}^2] = \frac{1}{(n-1)^2}(nC_4 + n(n-1)(\sigma^2+\mu^2)^2) = (\sigma^2+\mu^2)^2 + o(1).$$

Similarly,

$$\frac{2}{(n-1)^2 n} \sum_{i_1, i_2, i_3} \mathbb{E}[X_{i_1}^2 X_{i_2} X_{i_3}] = \frac{2}{(n-1)^2 n}(nC_4 + 2n(n-1)C_3\mu + n(n-1)\sigma^4 + (n^3 - 3n^2 + 2n)\sigma^2\mu^2)$$
$$= 2\sigma^2\mu^2 + o(1),$$

and

$$\frac{1}{n^4} \sum_{i_1, i_2, i_3, i_4} \mathbb{E}[X_{i_1} X_{i_2} X_{i_3} X_{i_4}] = \mu^4 + o(1).$$

To conclude, $\mathbb{E}[V_n(X)^2] \to \sigma^2 = \mathbb{E}[V_n(X)]^2$, so that the variance of $V_n(X)$ tends to zero.

3. First, the result is trivial if $\overline{X}_n$ and $V_n(X)$ are replaced by their limits, since for any n,
$$\frac{X_1 + \ldots + X_n - n\mu}{\sqrt{n\sigma^2}} \sim \mathcal{N}(0,1).$$
We now apply question 1:
$$\frac{X_1 + \ldots + X_n - n\mu}{\sqrt{n\sigma^2}} \Rightarrow \mathcal{N}(0,1) \qquad \text{and} \qquad \sqrt{V_n(X)} \Rightarrow \sqrt{\sigma^2}$$
implies
$$\frac{X_1 + \ldots + X_n - n\mu}{\sqrt{nV_n(X)}} = \frac{X_1 + \ldots + X_n - n\mu}{\sqrt{n\sigma^2}} \frac{\sqrt{V_n(X)}}{\sqrt{\sigma^2}} \Rightarrow \mathcal{N}(0,1).$$
If we replace μ with $\overline{X}_n$, the numerator vanishes identically, so that one has (trivially) convergence in distribution to 0.

□

Exercise 4. Return to Sheet 4, exercise 2: let μ be a probability measure on $\mathbb{N}$, described as usual by a sequence $(\mu_k)_{k \in \mathbb{N}}$. A virus spreads among the population; at time t, a number $N(t)$ of viruses are alive; between times t and $t+1$, each virus dies and creates new viruses in a random, independent and identical way: each virus creates k new viruses with probability μ_k. At time zero, we start with a deterministic number $N(0)$ of viruses.

Recall that, in sheet 4, we assumed that $R_0 := \sum_{k=0}^{+\infty} k\mu_k < 1$ and concluded that almost surely, the epidemic disappears in finite time. We suppose now that $R_0 = 1$ and $\mu_0 \neq 0$.

1. Show that the expected number of viruses is constant in time.
2. We denote by X an integer-valued random variable following μ. Using Jensen's inequality, prove that, for every $\alpha \in [0,1)$, one has $\mathbb{E}[\alpha^X] > \alpha$. Deduce that, for any $\alpha_0 \in [0,1]$, the sequence $(\alpha_n)_{n\in\mathbb{N}}$ defined by $\alpha_{n+1} = \mathbb{E}[\alpha_n^X]$ converges to 1.
3. We let $\alpha_0 = \mu_0$. Prove that

$$\mathbb{P}[N(t+1)=0] = \mathbb{E}[\mu_0^{N(t)}] = \mathbb{E}[\alpha_1^{N(t-1)}].$$

Continue by induction, and show that $N(t)$ converges in probability to zero as $t \to +\infty$. Does one have almost sure convergence? Convergence in L^1?

Solution of exercise 4.

1. From time t to $t+1$, one has $N(t)$ viruses that die; each virus j creates X_j new viruses; hence the family $(X_j)_{1\le j\le N(t)}$ is independent and identically distributed following μ. In particular, (using Exercise 1 of Sheet 3),

$$\mathbb{E}[N(t+1)] = \mathbb{E}\left[\sum_{j=1}^{N(t)} X_j\right] = \mathbb{E}[N(t)] \underbrace{\sum_{k=1}^{+\infty} k\mu_k}_{=1}.$$

2. Since $\mu_0 \neq 0$, with non-zero probability X is different from its mean $\mathbb{E}[X] = 1$. Since the function $t \mapsto \alpha^t = e^{t\log(\alpha)}$ is strictly convex on $\mathbb{R}$, one can apply the strict Jensen's inequality: $\mathbb{E}[\alpha^X] > \alpha^{\mathbb{E}[X]} = \alpha$.

   Let now $f : \alpha \to \mathbb{E}[\alpha^X]$. This function obviously maps $[0,1]$ into itself. Moreover $f(1) = 1$, and for any $\alpha < 1$ one has $f(\alpha) > \alpha$.

   In particular, the sequence α_n is non-decreasing and bounded from above (by 1) so it must converge to some limit ℓ. Since f is continuous by the dominated convergence theorem (one dominates by the function 1 whose expectation is $1 < \infty$), one has $f(\ell) = \ell$, so that $\ell = 1$.

3. One has $N(t+1) = 0$ when none of the $N(t)$ viruses alive at time t had any offspring. If there are k viruses alive, the probability that none of them have any offspring is μ_0^k, so that

$$\mathbb{P}[N(t+1)=0] = \mathbb{P}[N(t)=0] + \mathbb{P}[N(t)=1]\mu_0 + \mathbb{P}[N(t)=2]\mu_0^2 + \ldots = \mathbb{E}[\mu_0^{N(t)}].$$

   Now, as before, $N(t) = \sum_{j=1}^{N(t-1)} X_j$ where the X_j's are independent, identically distributed and follow μ. In particular, for any α,

$$\mathbb{E}[\alpha^{N(t)}] = \mathbb{E}\left[\prod_{j=1}^{N(t-1)} \alpha^{X_j}\right].$$

We now proceed similarly as in Exercise 1 of sheet 3; since the X_j's are independent form $N(t-1)$, this expectation is

$$\sum_{k=0}^{+\infty} \mathbb{P}[N(t-1)=k]\mathbb{E}[\alpha^X]^k = \mathbb{E}[f(\alpha)^{N(t-1)}].$$

In particular,

$$\mathbb{E}[\mu_0^{N(t)}] = \mathbb{E}[f(\mu_0)^{N(t-1)}] = \mathbb{E}[\alpha_1^{N-1}]$$

and, by induction,

$$\mathbb{E}[\mu_0^{N(t)}] = \mathbb{E}[\alpha_t^{N(0)}] = \alpha_t^{N(0)},$$

since $N(0)$ is constant (deterministic).

We can conclude: $\alpha_t \to 1$ as $t \to +\infty$, so $\mathbb{P}[N(t)=0] \to 1$ as $t \to +\infty$.

One has, in fact, almost sure convergence without having to apply Borel-Cantelli: the events $\{N(t)=0\}$ are increasing, so that

$$\mathbb{P}[\limsup_t (N(t)=0)] = \lim \mathbb{P}[N(t)=0] = 1,$$

where $\limsup_t (N(t)=0) =$ "the epidemic disappears in finite time".

However, $N(t)$ does not converge to 0 in L^1 since by question 1 $\mathbb{E}[N(t)] = \mathbb{E}[|N(t)|]$ is constant.

□