Law of the sum of Bernoulli random variables

Nicolas Chevallier
Université de Haute Alsace, 4, rue des frères Lumière 68093 Mulhouse
nicolas.chevallier@uha.fr

December 2006

Abstract

Let Δ_n be the set of all possible joint distributions of n Bernoulli random variables $X_1, ..., X_n$. Suppose that Δ_n which is a simplex in the 2^n-dimensional space, is endowed with the normalized Lebesgue measure μ_n. Suppose also that the integer n is large. Then we show that there is subset Δ of Δ_n whose measure $\mu_n(\Delta)$ is very close to 1, such that if the joint distribution of $(X_1, ..., X_n)$ is in Δ then the law of the sum $X_1 + ... + X_n$ is close to the binomial law $\mathcal{B}(n, \frac{1}{2})$. This result doesn't need any independance assumption. Next, we show a result of the same kind when Δ_n is endowed with an other probability measure ν_n.

Key words: Bernoulli random variable, Binomial law.
AMS Subject Classification 2000: 60-xx, 60Exx.

1 Introduction

The most common explanation for the ubiquity of the Gaussian law is the Central Limit Theorem. Another explanation closely related to the previous one, is that the Gaussian law is the only stable law with finite variance. The proofs of the Central Limit Theorem always rest on some independence assumption or at least on some stationarity assumption. The purpose of our work is to give in a very simplified situation, another kind of explanation for the ubiquity of the Gaussian law.

Consider a sequence $X_1, ..., X_n$ of Bernoulli random variables. We are interested in the law of the sum $S_n = X_1 + ... + X_n$ without any independance assumption about the random variables X_i. Since the Laplace-Moivre theorem asserts that up to a suitable normalization, for n large (and p not too small), the binomial distribution $\mathcal{B}(n, p)$ is close to the Gaussian law, an explanation for the ubiquity of the Gaussian law may be in our setting:
when n is large, the law of S_n is often very close to the symmetric binomial distribution $\mathcal{B}(n, \frac{1}{2})$.

We must explain what we mean by very often. Fix a positive integer n. Denote by Δ_n the set all possible joint distributions of n Bernoulli random variables $X_1, ..., X_n$. Δ_n is the set of all probability measures on $\{0, 1\}^n$. To each element $p = (p_i)_{i \in \{0,1\}^n}$ of Δ_n, one can associate the law of the sum $S_n = X_1 + ... + X_n$. It is a probability measure $L_n(p)$ on the set $\{0, ..., n\}$. When we choose the uniform probability law $b = (2^{-n}, ..., 2^{-n})$ on $\{0, 1\}^n$, we get $L_n(b) = \mathcal{B}_n = \mathcal{B}(n, \frac{1}{2})$ the symmetric binomial distribution:

$$\mathcal{B}_n(k) = C_n^k 2^{-n}$$

for all k in $\{0,...,n\}$. Let $\varepsilon > 0$, our aim is to estimate the size of the set $\Delta_{n,\varepsilon}$ of probability measures p in Δ_n such that for all k in $\{0,...,n\}$,

$$|L_n(p)(k) - \mathcal{B}_n(k)| \le \varepsilon.$$

To make precise this question, we have to measure the size of subsets of Δ_n. This can be done with the help of a probability measure on Δ_n. If μ is a probability measure on Δ_n, we wish to prove that

$$\mu(\Delta_{n,\varepsilon})$$

is close to 1 when n is large. In a less formal language :
when n is large, choosing at random the joint distribution of $(X_1,...,X_n)$, it is likely that the law of S_n is very close to the symmetric binomial distribution $\mathcal{B}(n,\frac{1}{2})$.

2 Statements of results

There are many choices for the probability measure μ and we shall only consider two. The first and the most natural one is $\mu = \mu_{1,n}$ the normalized Lebesgue measure on Δ_n. Our first result is:

Theorem 1 *There exists a constant A such that for all positive integers n, and all positive numbers ε,*

$$\mu_{1,n}(\Delta_{n,\varepsilon}) \ge 1 - \frac{A\sqrt{n}}{\varepsilon^2 2^{n-1}}$$

and

$$\mu_{1,n}(\{p \in \Delta_n : \sup_{I \subset \{0,...,n\}} |L_n(p)(I) - \mathcal{B}_n(I)| \le \varepsilon\}) \ge 1 - \frac{An^{5/2}}{\varepsilon^2 2^{n-1}}.$$

Though, the Lebesgue measure is very natural it has a drawback. One can consider the law of the first $n-1$ Bernoulli random variables $(X_1,...,X_{n-1})$ as a random variable defined on Δ_n. This random variable is the projection

$$\begin{aligned} \mathrm{pro}_n \quad &: \quad \Delta_n \longrightarrow \Delta_{n-1} \\ &: \quad (p_i)_{i\in\{0,1\}^n} \to (p_{(j,0)} + p_{(j,1)})_{j\in\{0,1\}^{n-1}} \end{aligned}$$

The point is that the Lebesgue measure $\mu_{1,n-1}$ is not the image of $\mu_{1,n}$ by the map pro_n; the family $(\mu_{1,n})_{n\ge1}$ is not a projective family of probability measures. We would like to find a projective family $(\mu_{2,n})_{n\ge1}$ of natural probability measures on the sequence of sets $(\Delta_n)_{n\ge1}$. This can be done inductively: when we know the law $(p_i)_{i\in\{0,1\}^n}$ of $(X_1,...,X_n)$ the law $(p'_{(i,j)})_{i\in\{0,1\}^n, j\in\{0,1\}}$ of $(X_1,...,X_n,X_{n+1})$ is chosen at random uniformly among all the possible laws. Let us make it precise. Consider the natural map which is almost a bijection,

$$\begin{aligned} \psi_n \quad &: \quad \Delta_n \times [0,1]^{\{0,1\}^n} \to \Delta_{n+1} \\ &: \quad ((p_i)_{i\in\{0,1\}^n}, (x_i)_{i\in\{0,1\}^n}) \to (p'_{(i,j)})_{i\in\{0,1\}^n,, j\in\{0,1\}} \end{aligned}$$

where $p'_{(i,0)} = p_i x_i$ and $p'_{(i,1)} = p_i(1-x_i)$. The uniformity means that the choice of $(p'_{(i,j)})_{i\in\{0,1\}^n,, j\in\{0,1\}}$ given $(p_i)_{i\in\{0,1\}^n}$, is done at random with respect to the Lebesgue measure λ_n on $[0,1]^{\{0,1\}^n}$. This enable to transfer a probability measure from Δ_n to Δ_{n+1}. If we have a probability measure $\mu_{2,n}$ on Δ_n, the product of this measure with the Lebesgue measure λ_n on $[0,1]^{\{0,1\}^n}$ gives rise to a measure on $\Delta_n \times [0,1]^{\{0,1\}^n}$ and its image by ψ_n is a new measure $\mu_{2,n+1}$ on Δ_{n+1}. Since the map $\mathrm{pro}_{n+1} \circ \psi_n : \Delta_n \times [0,1]^{\{0,1\}^n} \to \Delta_n$ is the projection on Δ_n, the image by pro_{n+1} of $\mu_{2,n+1}$ is $\mu_{2,n}$. Taking $\mu_{2,1}$ the normalized Lebesgue measure on Δ_1, we get a sequence $(\mu_{2,n})_{n\geq 1}$ of natural probability measures on the sequence of simplices $(\Delta_n)_{n\geq 1}$ (see section 3, for a purely probabilistic point of view about the measures $\mu_{2,n}$). This is our second choice which also has a drawback for it looses the symmetry between the random variables

$$\begin{aligned} \mathrm{pr}_i \quad &: \quad \Delta_n \to \mathbf{R} \\ &: \quad (p_j)_{j\in\{0,1\}^n} \to p_i, \end{aligned}$$

$i \in \{0,1\}^n$ as well as the symmetry between the variables X_i, $i \in \{1,...,n\}$. Our second result is:

Theorem 2 *There exists a constant C such that for all positive integers n, and all positive numbers ε,*

$$\mu_{2,n}(\Delta_{n,\varepsilon}) \geq 1 - \frac{C \ln n}{n^{3/2}\varepsilon^2}$$

and

$$\mu_{2,n}(\{p \in \Delta_n : \sup_{I\subset\{0,...,n\}} |L_n(p)(I) - \mathcal{B}_n(I)| \leq \varepsilon\}) \geq 1 - \frac{C \ln^{3/2} n}{\sqrt{n}\varepsilon^2}.$$

Remark. In both Theorems 1 and 2 it is possible to find an explicit value for the constants A and C. It is easy to check that the value $A = 2$ works in Theorem 1 whereas it is more difficult to give an explicit value for the constant C and we do not give any.

The main point of these two results is that they do not need any independence assumption about the variables X_i. There must be some other works of the same kind but we have only find one: K. Takeuchi and A. Takemura ([T,T]) have studied the law of the sum $S_n = X_1 + ... + X_n$ where the X_i are Bernoulli variables. They only assume some condition about "central binomial moments" which are a one to one function of the factorial moments (the k^{th} factorial moment of random variable X is $E(X(X-1)...(X-k+1))$). This allows them to prove convergence to the normal law or to the Poisson's law for a triangular array of Bernoulli variables $X_{i,n}$. Their hypothesis are only about the central binomial moments of $S_n = X_{1,n} + ... + X_{n,n}$.

3 Sketch of proofs

The ideas of the proofs of Theorems 1 and 2 are exactly the same. It is the reason why, although these proofs are not difficult, we begin by describing their main steps. For $i = 1, 2$, $n \in \mathbf{N}^*$, and $k \in \{0, ..., n\}$, denote by $E_{i,n}(k)$ and $V_{i,n}(k)$ the expectation and the variance of the random variable $p \to L_n(p)(k)$ defined on the probability space $(\Delta_n, \mu_{i,n})$.

Making use of the symmetries, we replace the simplex Δ_n by another geometrical space where the computation of expectations are easier. Then we show in both cases, that

$$E_{i,n}(k) = C_n^k 2^{-n} = \mathcal{B}_n(k).$$

Since by the Chebyshev inequality,

$$\mu_{i,n}(\{p \in \Delta_n : |L_n(p)(k) - E_{i,n}(k)| > \varepsilon\}) \leq \frac{V_{i,n}(k)}{\varepsilon^2},$$

Theorems 1 and 2 can be deduced from appropriates upper bound on the variances $V_{i,n}(k)$. In the first case, standard results lead to the inequality

$$V_{1,n}(k) \le \frac{C_n^k}{2^{2n}}.$$

In the second case, computations are not as easy as in the first case. Some well known estimates about binomial coefficients, enable to show that for all k in $\{0, ..., n\}$,

$$V_{2,n}(k) \le \frac{C}{n^2},$$

and that for all k such that $\left|k - \frac{n}{2}\right| \ge 2\sqrt{n \ln n}$,

$$V_{2,n}(k) \le \frac{C}{n^5},$$

where C is a constant independent of n and k.

4 Proof of theorem 1

1. The cardinal number of the set Δ_n is $N = 2^n$ and there is a one to one correspondence between the sets $\{0,1\}^n$ and $\{0, ..., N-1\}$. Therefore each p in Δ_n can be seen as a probability measure on $\{0, 1, ..., N-1\}$:

$$\Delta_n = \{(p_0, ..., p_{N-1}) \in \mathbf{R}^N : p_0, p_2, ..., p_{N-1} \ge 0,\ p_0 + p_2 + ... + p_{N-1} = 1\}.$$

Furthermore, for all k in $\{0, ..., n\}$, there is a subset E_k of $\{0, ..., N-1\}$ with C_n^k elements, such that for all $p = (p_0, ..., p_{N-1})$ in Δ_n,

$$L_n(p)(k) = \sum_{i \in E_k} p_i.$$

2. Let σ be a permutation of the set $\{0, ..., N-1\}$. σ induces the linear map, $f_\sigma(x_1, ..., x_n) = (x_{\sigma(0)}, x_{\sigma(2)}, ..., x_{\sigma(N-1)})$ which sends Δ_n onto itself. Therefore the measure $\mu_{1,n}$ is f_σ-invariant. It follows that given a subset E of $\{0, ..., N-1\}$, the distribution function of the map

$$p = (p_0, ..., p_{N-1}) \in \Delta_n \to L_E(p) = \sum_{i \in E} p_i$$

depends only on the cardinal number of E. This means that for all $k \in \{0, ..., n\}$, the map L_{E_k} have the same distribution than the map L_{F_k} where

$$F_k = \{0, ..., C_n^k - 1\}.$$

Hence,

$$\mu_{1,n}(\{p \in \Delta_n : |L_n(p)(k) - \mathcal{B}_n(k)| \ge \varepsilon\}) = \mu_{1,n}(\{p \in \Delta_n : |L_{F_k}(p) - \mathcal{B}_n(k)| \ge \varepsilon\}).$$

3. In order to estimate $\mu_{1,n}(\{p \in \Delta_n : |L_{F_k}(p) - \mathcal{B}_n(k)| \ge \varepsilon\})$, let us introduce another way to see the probability space $(\Delta_n, \mu_{1,n})$. Let $(Y_1, ..., Y_{N-1})$ be $N-1$ independent random variables uniformly distributed in the interval $[0,1]$. Arranging them in ascending order we find $N-1$ random variables $Z_1 \le Z_2 \le ... \le Z_{N-1}$. The joint distribution of $(Z_1, ..., Z_{N-1})$ is the normalized Lebesgue measure ν on

$$T = \{(z_1, ..., z_{N-1}) \in \mathbf{R}^{N-1} : 0 \le z_1 \le z_2 \le ... \le z_{N-1} \le 1\}.$$

Let $\phi : \mathbf{R}^{N-1} \to \mathbf{R}^N$ be the map defined by

$$\phi(z_1, ..., z_{N-1}) = (z_1, z_2 - z_1, ..., z_{N-1} - z_{N-2}, 1 - z_{N-1}).$$

The image of T by ϕ is Δ_n and since the map ϕ is affine, the image of the measure ν by ϕ is the measure $\mu_{1,n}$. Now, let $F = \{0, ..., m-1\}$ be a subset of $\{0, ..., N-1\}$. For all $z = (z_1, ..., z_{N-1}) \in T$, we have

$$L_F(\phi(z)) = z_1 + (z_2 - z_1) + ... + (z_m - z_{m-1}) = z_m,$$

hence the distribution of L_F is the same as the distribution of the map

$$R_m : (z_1, ..., z_{N-1}) \in T \to z_m.$$

It follows that

$$E_{\mu_{1,n}}(L_F) = E_\nu(R_m)$$

and the same holds for the variances

$$V_{\mu_{1,n}}(L_F) = V_\nu(R_m).$$

4. The distribution of R_m is well known, its density h_m is given by the formula

$$h_m(t) = \frac{(N-1)!}{(m-1)!(N-1-m)!} \times t^{m-1}(1-t)^{N-1-m}$$

(see [Da, Du]). Therefore,

$$\begin{aligned} E_\nu(R_m) &= \int_0^1 t \times \frac{(N-1)!}{(m-1)!(N-1-m)!} t^{m-1}(1-t)^{N-1-m} dt \\ &= \frac{(N-1)!}{(m-1)!(N-1-m)!} \times \frac{\Gamma(m+1)\Gamma(N-m)}{\Gamma(N+1)} \\ &= \frac{(N-1)!}{(m-1)!(N-1-m)!} \times \frac{m!(N-m-1)!}{N!} \\ &= \frac{m}{N}, \end{aligned}$$

$$\begin{aligned} E_\nu(R_m^2) &= \int_0^1 t^2 \times \frac{(N-1)!}{(m-1)!(N-1-m)!} t^{m-1}(1-t)^{N-1-m} dt \\ &= \frac{(N-1)!}{(m-1)!(N-1-m)!} \times \frac{\Gamma(m+2)\Gamma(N-m)}{\Gamma(N+2)} \\ &= \frac{(N-1)!}{(m-1)!(N-1-m)!} \times \frac{(m+1)!(N-m-1)!}{(N+1)!} \\ &= \frac{m \times (m+1)}{N \times (N+1)} \end{aligned}$$

and

$$\begin{aligned} V_\nu(R_m) = E_\nu(R_m^2) - E_\nu(R_m)^2 &= \frac{m(m+1)}{N(N+1)} - \frac{m^2}{N^2} = \frac{Nm(m+1) - (N+1)m^2}{N^2(N+1)} \\ &= \frac{Nm^2 + Nm - Nm^2 - m^2}{N^2(N+1)} = \frac{m(N-m)}{N^2(N+1)} \leq \frac{m}{N^2}. \end{aligned}$$

Coming back to L_{E_k}, we find that for all k in $\{0, ..., n\}$,

$$\begin{aligned} E_{1,n}(k) &= E_{\mu_{1,n}}(L_{E_k}) = \frac{C_n^k}{N} = C_n^k 2^{-n} = \mathcal{B}_n(k), \\ V_{1,n}(k) &= V_{\mu_{1,n}}(L_{E_k}) \leq \frac{C_n^k}{N^2}. \end{aligned}$$

Making use of the Stirling formula, it is easy to see that $C_n^k \le A\frac{2^n}{\sqrt{n}}$ where A is a constant independent of n. It follows that $V_{1,n}(k) \le \frac{A}{2^n\sqrt{n}}$. Finally, with Chebyshev inequality we get

$$\mu_{1,n}(\{p \in \Delta_n : \left|L_{E_k}(p) - C_n^k 2^{-n}\right| \ge \varepsilon\}) \le \frac{A}{\varepsilon^2 2^n \sqrt{n}}$$

and

$$\mu_{1,n}(\{p \in \Delta_n : \max_{k=0}^{n} \left|L_{E_k}(p) - C_n^k 2^{-n}\right| \ge \varepsilon\}) \le (n+1) \times \frac{A}{\varepsilon^2 2^n \sqrt{n}},$$

therefore

$$\mu_{1,n}(\Delta_\varepsilon) \ge 1 - \frac{A\sqrt{n}}{\varepsilon^2 2^{n-1}}.$$

The second inequality of Theorem 1 follows from the first in replacing ε by ε/n. □

If we want to find an explicit value for A, we can use the following inequalities instead of the Stirling formula

$$\sqrt{2\pi n} n^n \exp(-n + \frac{1}{12n+1}) \le n! \le \sqrt{2\pi n} n^n \exp(-n + \frac{1}{12n}),$$

(see [Fe], p. 50-54). An easy calculation shows that the value $A = 2$ works.

5 About the definition of $\mu_{2,n}$

In this section we give two other ways to introduce the measure $\mu_{2,n}$: proposition 1 and 2. While proposition 1 is not needed for the following, proposition 2 is useful. It replace the simplices Δ_n, $n \in \mathbf{N}^*$, by a single product space endowed with a product probability.

Notations. 1. For x in $[0,1]$, we put $x^{(0)} = x$ and $x^{(1)} = 1 - x$.
2. Denote $\mathrm{pro}_n : \Delta_n \to \Delta_{n-1}$ the map defined by $\mathrm{pro}_n((p_i)_{i\in\{0,1\}^n}) = (p'_i)_{i\in\{0,1\}^{n-1}}$ where $p'_i = p_{i0} + p_{i1}$ for all $i \in \{0,1\}^{n-1}$.
3. For each i in $\{0,1\}^n$, denote $\mathrm{pr}_i : \mathbf{R}^{\{0,1\}^n} \to \mathbf{R}$ the map pr defined by $\mathrm{pr}((p_j)_{j\in\{0,1\}^n}) = p_i$. It induces a random variable on Δ_n and it is readily seen that for all i in $\{0,1\}^{n-1}$, $\mathrm{pr}_{i0} + \mathrm{pr}_{i1} = \mathrm{pr}_i \circ \mathrm{pro}_n$ on Δ_n.
4. For an integer $n \ge 2$, the map ψ_{n-1} is defined by:

$$\begin{array}{lll} \psi_{n-1} & : & \Delta_{n-1} \times [0,1]^{\{0,1\}^{n-1}} \to \Delta_n \\ & : & ((p_i)_{i\in\{0,1\}^{n-1}}, (x_i)_{i\in\{0,1\}^{n-1}}) \to (p'_{ij})_{i\in\{0,1\}^{n-1}, j\in\{0,1\}} \end{array}$$

where $p'_{i0} = p_i x_i$ and $p'_{i1} = p_i(1 - x_i)$.

In the introduction we present a rather geometric point of view about the probability measures $\mu_{2,n}$. It is possible to give a more probabilistic point of view about these probability measures. The simplex Δ_n is the set of all probability laws of a sequence $(X_1, ..., X_n)$ of n Bernoulli random variables and the map which associated to each law of $(X_1, ..., X_n)$ the law of the first $n-1$ Bernoulli random variables $(X_1, ..., X_{n-1})$ is just the projection $\mathrm{pro}_n : \Delta_n \to \Delta_{n-1}$. When we know the law of $(X_1, ..., X_{n-1})$ what can we expect about the law of the whole sequence $(X_1, ..., X_n)$? This is given by the conditional distribution given pro_n of the random variables

$$\frac{P(X_1 = i_1, ..., X_{n-1} = i_{n-1}, X_n = 0)}{P(X_1 = i_1, ..., X_{n-1} = i_{n-1})} = \frac{\mathrm{pr}_{(i,0)}}{\mathrm{pr}_i \circ \mathrm{pro}_n} : \Delta_n \to [0,1],$$

$i = (i_1, ...i_{n-1}) \in \{0,1\}^{n-1}$. The probability measures $\mu_{2,n}$ are the only such that these variables are all uniformly distributed in the interval $[0,1]$ and independent conditionally to the law of $(X_1, ..., X_{n-1})$. This the meaning of the next proposition which we state without proof.

Proposition 1 *The sequence $(\mu_{2,n})_{n\geq 1}$ is the unique sequence of probability measures such that :*
i. $\mu_{2,1}$ is the normalized Lebesgue measure on Δ_1,
ii. for all integer $n \geq 1$, $\mu_{2,n}$ is a probability on Δ_n,
iii. for all integer $n \geq 2$, the image of $\mu_{2,n}$ by pro_n is $\mu_{2,n-1}$,
iv. for all integer $n \geq 2$ and all family $(B_i)_{i\in\{0,1\}^{n-1}}$ of Borel subsets of $[0,1]$,

$$\mu_{2,n}(\frac{\mathrm{pr}_{(i,0)}}{\mathrm{pr}_i \circ \mathrm{pro}_n} \in B_i,\ i \in \{0,1\}^{n-1} | \mathrm{pro}_n) = \prod_{i\in\{0,1\}^{n-1}} \lambda(B_i)$$

where λ is the Lebesgue measure on $[0,1]$.

Remark. iv can be replace as well by:
for all integers $n \geq 2$ and all family $(B_i)_{i\in\{0,1\}^{n-1}}$ of Borel subsets of $[0,1]$,

$$\mu_{2,n}(\mathrm{pr}_{(i,0)} \in B_i,\ i \in \{0,1\}^{n-1} | \mathrm{pro}_n) = \prod_{i\in\{0,1\}^{n-1}} \frac{\lambda(B_i \cap]0, \mathrm{pr}_i[)}{\mathrm{pr}_i}$$

where λ is the Lebesgue measure on $[0,1]$.

By definition, the probability $\mu_{2,n}$ is the image by ψ_{n-1} of the probability $\mu_{2,n-1} \otimes \lambda_{n-1}$ where λ_{n-1} is the Lebesgue measure on $[0,1]^{\{0,1\}^{n-1}}$. We can iterate this process from n down to 1 and we see that the probability measure $\mu_{2,n}$ is the image of the Lebesgue measure on

$$\Omega_n = [0,1]^{\{\emptyset\}} \times [0,1]^{\{0,1\}} \times ... \times [0,1]^{\{0,1\}^{n-1}}$$

by a map $\phi_n : \Omega_n \to \Delta_n$. It will be more efficient to define ϕ_n on an unique probability space Ω which does not depend on n.

Notations. Denote by $\mathcal{J}$ the set $\{\emptyset\} \cup (\cup_{k=1}^{\infty} \{0,1\}^k)$ and Ω the set $[0,1]^{\mathcal{J}}$. For $j = (j_1, ..., j_k)$ in $\mathcal{J}$, denote by $Z_j : \Omega = [0,1]^{\mathcal{J}} \to [0,1]$ the random variable defined by $Z_j((\omega_i)_{i\in\mathcal{J}}) = \omega_j$.
Denote by Q the infinite product of Lebesgue measures on Ω.

Proposition 2 *For all integers $n \geq 1$, consider the map $\phi_n : \Omega \to \mathbf{R}^{\{0,1\}^n}$ defined by*

$$\mathrm{pr}_{(i_1,...,i_n)} \circ \phi_n = Z_{\emptyset}^{(i_1)} Z_{(i_1)}^{(i_2)} Z_{(i_1,i_2)}^{(i_3)} ... Z_{(i_1,...,i_{n-1})}^{(i_n)}.$$

Then $\phi_n(\Omega) \subset \Delta_n$ and the image by ϕ_n of the probability measure Q is $\mu_{2,n}$.

Proof. It is easy to check that $\phi_n(\omega)$ is in Δ_n for all ω in Ω_n. Indeed,

$$\mathrm{pr}_{(i_1,...,i_n)} \circ \phi_n = \mathrm{pr}_{(i_1,...,i_{n-1})} \circ \phi_{n-1} \times Z_{(i_1,...,i_{n-1})}^{(i_n)},$$

thus

$$\begin{aligned} \mathrm{pr}_{(i_1,...i_{n-1},0)}(\phi_n(\omega)) + \mathrm{pr}_{(i_1,...,i_{n-1},1)}(\phi_n(\omega)) &= \mathrm{pr}_{(i_1,...,i_{n-1})}(\phi_{n-1}(\omega)) \times Z_{(i_1,..,i_{n-1})} \\ &+ \mathrm{pr}_{(i_1,...,i_{n-1})}(\phi_{n-1}(\omega)) \times (1 - Z_{(i_1,..,i_{n-1})}) \\ &= \mathrm{pr}_{(i_1,...,i_{n-1})}(\phi_{n-1}(\omega)) \end{aligned}$$

and it follows by induction that

$$\sum_{i\in\{0,1\}^n} \mathrm{pr}_i(\phi_n(\omega)) = \sum_{i\in\{0,1\}} \mathrm{pr}_i(\phi_1(\omega)) = \mathrm{pr}_0(\phi_1(\omega))+\mathrm{pr}_1(\phi_1(\omega)) = Z_\emptyset^{(0)}+Z_\emptyset^{(1)} = 1.$$

Next we prove by induction that the image by ϕ_n of the probability measure Q, is $\mu_{2,n}$. Suppose the image by ϕ_{n-1} of the probability measure Q is $\mu_{2,n-1}$. Using the sequence of maps $(\psi_n)_{n\geq 1}$, it is easy to find an induction relation satisfied by the sequence of maps $(\phi_n)_{n\geq 1}$, we have

$$\phi_n(\omega) = \psi_{n-1}(\phi_{n-1}(\omega), (Z_i(\omega))_{i\in\{0,1\}^{n-1}}).$$

Now ϕ_{n-1} and $(Z_i)_{i\in\{0,1\}^{n-1}}$ are independent random variables, therefore the image by $\omega \in \Omega \to (\phi_{n-1}(\omega), (Z_i(\omega))_{i\in\{0,1\}^{n-1}})$ of the probability measure Q is the product of the image by ϕ_{n-1} of Q and of the Lebesgue measure on $[0,1]^{\{0,1\}^{n-1}}$ which is λ_{n-1}. By induction hypothesis we get $\mu_{2,n-1}\otimes\lambda_{n-1}$ and by definition, the image of $\mu_{2,n-1}\otimes\lambda_{n-1}$ by ψ_{n-1} is $\mu_{2,n}$. $\square$

6 Calculation of the first two moments of $p \in \Delta_n \to L_n(p)(k)$

Notations. Let n be a positive integer.
1. For each subset F of $\{0,1\}^n$, we shall denote by L_F the map defined by

$$p = (p_i)_{i\in\{0,1\}^n} \in \Delta_n \to L_F(p) = \sum_{i\in F} p_i.$$

2. For all integers k in $\{0,...n\}$, $F_{n,k}$ denote the subset of element $i = (i_l)_{l\in\{1,...,n\}} \in \{0,1\}^n$ such that $\sum_{l=1}^n i_l = k$.

Let $n \geq 1$ be an integer and let k be an integer in $\{0,...,n\}$. We would like to estimate

$$\begin{aligned} E_{2,n}(k) = E_{\mu_{2,n}}(L_n(.)(k)) &= E_{\mu_{2,n}}(L_{F_{n,k}}) \\ &= \int_{\Delta_n} \sum_{i\in F_k} p_i \, d\mu_{2,n}((p_i)_{i\in\{0,1\}^{n-1}}) \\ &= \int_\Omega \sum_{i\in F_k} \mathrm{pr}_i(\phi_n(\omega)) \, dQ(\omega). \end{aligned}$$

and

$$\begin{aligned} E_{\mu_{2,n}}(L_n^2(.)(k)) = E_{\mu_{2,n}}(L_{F_{n,k}}^2) &= \int_{\Delta_n} (\sum_{i\in F_k} p_i)^2 \, d\mu_{2,n}((p_i)_{i\in\{0,1\}^{n-1}}) \\ &= \int_\Omega (\sum_{i\in F_k} \mathrm{pr}_i(\phi_n(\omega)))^2 \, dQ(\omega). \end{aligned}$$

Set $f_{n,k} = \sum_{i\in F_k} \mathrm{pr}_i \circ \phi_n$.

1. By proposition 2,

$$\begin{aligned} E_Q(f_{n,k}) = E_Q(\sum_{i\in F_{n,k}} \mathrm{pr}_i \circ \phi_n) &= E_Q(\sum_{(i_1,\dots,i_n)\in F_{n,k}} \mathrm{pr}_{(i_1,\dots,i_{n-1})} \circ \phi_{n-1} \times Z_{(i_1,\dots,i_{n-1})}^{i_n}) \\ &= \sum_{(i_1,\dots,i_n)\in F_{n,k}} E_Q(\mathrm{pr}_{(i_1,\dots,i_{n-1})} \circ \phi_{n-1} \times Z_{(i_1,\dots,i_{n-1})}^{(i_n)})) \end{aligned}$$

and since ϕ_{n-1} and $Z^{i_n}_{(i_1,...,i_{n-1})}$ are independent,

$$E_Q(f_{n,k}) = \sum_{(i_1,...,i_n)\in F_{n,k}} E_Q(\text{pr}_{(i_1,...,i_{n-1})} \circ \phi_{n-1}) E_Q(Z^{(i_n)}_{(i_1,...,i_{n-1})}))$$
$$= \frac{1}{2} \sum_{(i_1,...,i_n)\in F_{n,k}} E_Q(\text{pr}_{(i_1,...,i_{n-1})} \circ \phi_{n-1}).$$

Furthermore, $F_{n,k} = F_{n-1,k-1} \times \{1\} \cup F_{n-1,k} \times \{0\}$, thus

$$E_Q(f_{n,k}) = \frac{1}{2}(\sum_{(i_1,...,i_n)\in F_{n-1,k-1}\times\{1\}} E_Q(\text{pr}_{(i_1,...,i_{n-1})} \circ \phi_{n-1}) + \sum_{(i_1,...,i_n)\in F_{n-1,k}\times\{0\}} E_Q(\text{pr}_{(i_1,...,i_{n-1})} \circ \phi_{n-1}))$$
$$= \frac{1}{2}(E_Q(f_{n-1,k-1}) + E_Q(f_{n-1,k})).$$

We have also

$$E_Q(f_{1,0}) = E_Q(Z^{(0)}_\emptyset) = E_Q(Z^{(1)}_\emptyset) = E_Q(f_{1,1}) = \frac{1}{2},$$

therefore, by induction, we get

$$E_Q(f_{n,k}) = C^k_n 2^{-n}.$$

Hence

$$E_{2,n}(k) = C^k_n 2^{-n}.$$

2. The quadratic mean $E_Q(f^2_{n,k})$ is a little more difficult to estimate. The main idea is to decompose $F_{n,k}$ in the two sets

$$F^0_{n,k} = \{(i_1, ..., i_n) \in F_{n,k} : i_1 = 0\},$$
$$F^1_{n,k} = \{(i_1, ..., i_n) \in F_{n,k} : i_1 = 1\}$$

and to observe that for each i in $F^0_{n,k}$ and each j in $F^1_{n,k}$, the two variables

$$Z^{(i_2)}_{(i_1)} Z^{(i_3)}_{(i_1,i_2)} ... Z^{(i_n)}_{(i_1,...,i_{n-1})}, \quad Z^{(j_2)}_{(j_1)} Z^{(j_3)}_{(j_1,j_2)} ... Z^{(j_n)}_{(j_1,...,j_{n-1})}$$

are independent. We have

$$E_Q(f^2_{n,k}) = E_Q((\sum_{i\in F^0_{n,k}} \text{pr}_i \circ \phi_n)^2) + E_Q((\sum_{i\in F^1_{n,k}} \text{pr}_i \circ \phi_n)^2)$$
$$+ 2E_Q((\sum_{i\in F^0_{n,k}} \text{pr}_i \circ \phi_n)(\sum_{j\in F^1_{n,k}} \text{pr}_i \circ \phi_n))$$
$$= T_1 + T_2 + 2T_3.$$

The first term gives

$$T_1 = E_Q((\sum_{i\in F^0_{n,k}} Z^{(0)}_\emptyset Z^{(i_2)}_{(0)} Z^{(i_3)}_{(0,i_2)} ... Z^{(i_n)}_{(0,i_2...,i_{n-1})})^2)$$
$$= E_Q(Z^2_\emptyset(\sum_{i\in F^0_{n,k}} Z^{(i_2)}_{(0)} Z^{(i_3)}_{(0,i_2)} ... Z^{(i_n)}_{(0,i_2...,i_{n-1})})^2),$$

since $Z_\emptyset$ is independent of the others Z_i, we get

$$E_Q((\sum_{i\in F^0_{n,k}} \text{pr}_i \circ \phi_n)^2) = E_Q(Z^2_\emptyset) E_Q(\sum_{i\in F^0_{n,k}} Z^{(i_2)}_{(0)} Z^{(i_3)}_{(0,i_2)} ... Z^{(i_n)}_{(0,i_2,...,i_{n-1})})^2).$$

The last thing to see for the computation of the first term is that

$$E_Q((\sum_{i\in F^0_{n,k}} Z^{(i_2)}_{(0)} Z^{(i_3)}_{(0,i_2)}...Z^{(i_n)}_{(0,i_2,...,i_{n-1})})^2) = E_Q((\sum_{i\in F_{n-1,k}} Z^{(i_1)}_{(\emptyset)} Z^{(i_2)}_{(i_1)}...Z^{(i_{n-1})}_{(i_1,...,i_{n-2})})^2) = E_Q(f^2_{n-1,k}),$$

thus

$$T_1 = \frac{1}{3}E_Q(f^2_{n-1,k}).$$

Exactly the same arguments show that

$$T_2 = \frac{1}{3}E_Q(f^2_{n-1,k-1}).$$

By independence, the last term gives

$$\begin{aligned}
T_3 &= E_Q((\sum_{i\in F^0_{n,k}} Z^{(0)}_{\emptyset} Z^{(i_2)}_{(0)} Z^{(i_3)}_{(0,i_2)}...Z^{(i_n)}_{(0,i_2,...,i_{n-1})})(\sum_{i\in F^1_{n,k}} Z^{(1)}_{\emptyset} Z^{(i_2)}_{(1)} Z^{(i_3)}_{(1,i_2)}...Z^{(i_n)}_{(1,i_2,...,i_{n-1})}))\\
&= E_Q(Z^{(0)}_{\emptyset} Z^{(1)}_{\emptyset})E_Q((\sum_{i\in F^0_{n,k}} Z^{(i_2)}_{(0)} Z^{(i_3)}_{(0,i_2)}...Z^{(i_n)}_{(0,i_2,...,i_{n-1})}))E_Q((\sum_{i\in F^1_{n,k}} Z^{(i_2)}_{(1)} Z^{(i_3)}_{(1,i_2)}...Z^{(i_n)}_{(0,i_2,...,i_{n-1})}))\\
&= \frac{1}{6}E_Q(f_{n-1,k})E_Q(f_{n-1,k-1}).
\end{aligned}$$

Finally, we get the relation

$$E_Q(f^2_{n,k}) = \frac{1}{3}[E_Q(f^2_{n-1,k}) + E_Q(f^2_{n-1,k-1}) + E_Q(f_{n-1,k})E_Q(f_{n-1,k-1})].$$

3. This recursion relation and the equality $E_Q(f_{m,l}) = 2^{-m}C^l_m$, enable to find a recursion relation between $V_Q(f_{n-1,k-1})$, $V_Q(f_{n-1,k})$ and $V_Q(f_{n,k})$:

$$\begin{aligned}
V_Q(f_{n,k}) &= E_Q(f^2_{n,k}) - E_Q(f_{n,k})^2\\
&= \frac{1}{3}[E_Q(f^2_{n-1,k}) + E_Q(f^2_{n-1,k-1}) + E_Q(f_{n-1,k})E_Q(f_{n-1,k-1})]\\
&- \frac{1}{4}[E_Q(f_{n-1,k}) + E_Q(f_{n-1,k-1})]^2\\
&= \frac{1}{3}(V_Q(f_{n-1,k}) + V_Q(f_{n-1,k-1}) + E_Q(f_{n-1,k})E_Q(f_{n-1,k-1}))\\
&+ \frac{1}{12}[E_Q(f_{n-1,k})^2 + E_Q(f_{n-1,k-1})^2] - \frac{1}{2}E_Q(f_{n-1,k})E_Q(f_{n-1,k-1})\\
&= \frac{1}{3}(V_Q(f_{n-1,k}) + V_Q(f_{n-1,k-1})) + \frac{1}{12}(E_Q(f_{n-1,k}) - E_Q(f_{n-1,k-1}))^2\\
&= \frac{1}{3}(V_Q(f_{n-1,k}) + V_Q(f_{n-1,k-1})) + \frac{1}{12}[2^{-n+1}(C^k_{n-1} - C^{k-1}_{n-1})]^2.
\end{aligned}$$

Hence,

$$V_{\mu_{2,n}}(L_{F_{n,k}}) = \frac{1}{3}[V_{\mu_{2,n-1}}(L_{F_{n-1,k}})+V_{\mu_{2,n-1}}(L_{F_{n-1,k-1}})]+\frac{1}{12}[2^{-n+1}(C^{k-1}_{n-1}-C^k_{n-1})]^2.$$

6.1 An upper bound for $V_{2,n}(k) = V_{\mu_{2,n}}(L_{F_{n,k}})$

We shall need the following lemma.

Lemma 1 *There exists a constant C such that for all integers $n \geq 1$ we have:*
1. for all k in $\{0, ..., n\}$,

$$2^{-n}|C^{k-1}_n - C^k_n| \leq \frac{c}{n}.$$

2. for all k in $\{0, ..., n\}$ such that $\left|k - \frac{n}{2}\right| \geq \sqrt{n \ln n}$,

$$2^{-n}|C_n^{k-1} - C_n^k| \leq \frac{c}{n^{5/2}}.$$

Proof. In what follows, C denotes a constant whose value may change at each line. Since

$$\begin{aligned} 2^{-n}|C_n^{k-1} - C_n^k| &= 2^{-n}\frac{n!}{(n-k)!k!}\left|\frac{k}{n-k+1} - 1\right| \\ &= 2^{-n}C_n^k\left|\frac{n+1-2k}{n+1-k}\right|, \end{aligned}$$

we can use the classical Laplace-Moivre estimate about the binomial law:
Let $(a_n)_{n\geq 1}$ be a sequence of non negative real numbers which go to 0 as n goes to infinity. Then for all positive integers n and all integers k such that $\left|k - \frac{n}{2}\right| \leq a_n n^{2/3}$ we have

$$2^{-n}C_n^k = \frac{1+\delta_n(k)}{\sqrt{\frac{\pi}{2}n}}\exp\left(-\frac{2(k-\frac{n}{2})^2}{n}\right)$$

where

$$\lim_{n\to\infty}\sup_{k:|k-n/2|\leq a_n n^{2/3}}|\delta_n(k)| = 0$$

(actually, it is a slight extension of the Laplace-Moivre theorem which deals only with integers k such that $\left|k - \frac{n}{2}\right| \leq a\sqrt{n}$ where a is a fix real number; see [Fe] p. 185 theorem 1, or [Le] p. 36 proposition 8.2). It follows that for all positive integers n and all integers k such that $\left|k - \frac{n}{2}\right| \leq a_n n^{2/3}$,

$$2^{-n}C_n^k \leq \frac{C}{\sqrt{n}}\exp\left(-\frac{2(k-\frac{n}{2})^2}{n}\right)$$

where the constant C does not depend on n. Making use of the monotonicity of the binomial coefficients, we get the following inequality

$$2^{-n}C_n^k \leq \frac{C}{n^{5/2}}$$

for all integers n and all integers k such that $\left|k - \frac{n}{2}\right| \geq \sqrt{n \ln n}$. This last inequality implies 2. Now let us prove 1. For all positive integers n and all integers k such that $\left|k - \frac{n}{2}\right| \leq \sqrt{n \ln n}$, we have

$$2^{-n}|C_n^{k-1} - C_n^k| \leq \frac{C}{\sqrt{n}}\exp\left(-\frac{2(k-\frac{n}{2})^2}{n}\right)\frac{|n-2k|+1}{n}.$$

Put $t = \frac{k-\frac{n}{2}}{\sqrt{n}}$. We get

$$2^{-n}|C_n^{k-1} - C_n^k| \leq \frac{C\,|t|}{n}\exp\left(-2t^2\right) + \frac{C}{n^{3/2}}$$

and since the function $t \in \mathbf{R} \to |t|\,e^{-2t^2}$ is bounded, $2^{-n}|C_n^{k-1} - C_n^k| \leq \frac{C}{n}$. □

Proposition 3 *There exists a constant C such that or all positive integers n,*

$$v_n := \sup_{k\in\{0,\ldots,n\}} V_{\mu_{2,n}}(L_{F_{n,k}}) \leq \frac{C}{n^2}$$

Proof. Let $n \geq 2$ be an integer. Since for all k in $\{0, ..., n\}$,

$$V_{\mu_{2,n}}(L_{F_{n,k}}) = \frac{1}{3}(V_{\mu_{2,n-1}}(L_{F_{n-1,k-1}})+V_{\mu_{2,n-1}}(L_{F_{n-1,k}}))+\frac{1}{12}[2^{-n+1}(C_{n-1}^{k-1}-C_{n-1}^{k})]^2,$$

we have

$$v_n \leq \frac{2}{3}v_{n-1} + \frac{1}{12}[2^{-n+1}(C_{n-1}^{k-1} - C_{n-1}^{k})]^2.$$

With the previous lemma we get

$$v_n \leq \frac{2}{3}v_{n-1} + \frac{C}{n^2}.$$

By induction, we get that for all integers $n \geq 2$,

$$v_n \leq (\frac{2}{3})^{n-1}v_1 + C\sum_{i=0}^{n-2} \frac{1}{(n-i)^2}(\frac{2}{3})^i.$$

The sum $\sum_{i=0}^{n-2} \frac{1}{(n-i)^2}(\frac{2}{3})^k$ is easy to estimate:

$$\sum_{i=0}^{n-2} \frac{1}{(n-i)^2}(\frac{2}{3})^i = \sum_{0\leq i\leq n/2} + \sum_{n/2<i\leq n-2},$$

since

$$\sum_{0\leq i\leq n/2} \frac{1}{(n-i)^2}(\frac{2}{3})^i \leq \frac{4}{n^2}\sum_{0\leq i\leq n/2}(\frac{2}{3})^i \leq \frac{12}{n^2}$$

and since

$$\sum_{n/2<i\leq n-2} \frac{1}{(n-i)^2}(\frac{2}{3})^i \leq \frac{n}{2}(\frac{2}{3})^{n/2} \leq \frac{C}{n^2},$$

we have

$$\begin{aligned} v_n &\leq (\frac{2}{3})^n v_1 + \frac{C}{n^2} \\ &\leq \frac{C}{n^2}. \square \end{aligned}$$

Proposition 4 *There exists a constant C such that or all positive integers n*

$$u_n := \sup_{k\in\{0,\dots,n\}:\left|k-\frac{n}{2}\right|\geq 2\sqrt{n\ln n}} V_{\mu_{2,n}}(L_{F_{n,k}}) \leq \frac{C}{n^5}.$$

Dem. Let $n \geq 2$ be an integer. First note that the two variables $L_{F_{n,k}}$ and $L_{F_{n,n-k}}$ have the same law, so it suffices to prove the proposition for $k \geq n/2$. By lemma 1, for all k in $\{0, ..., n\}$ such that $k - \frac{n}{2} \geq \sqrt{n\ln n}$, we have

$$\begin{aligned} V_{\mu_{2,n}}(L_{F_{n,k}}) &= \frac{1}{3}[V_{\mu_{2,n-1}}(L_{F_{n-1,k-1}}) + V_{\mu_{2,n-1}}(L_{F_{n-1,k}})] + \frac{1}{12}[2^{-n+1}(C_{n-1}^{k-1} - C_{n-1}^{k})]^2 \\ &\leq \frac{1}{3}[V_{\mu_{2,n-1}}(L_{F_{n-1,k-1}}) + V_{\mu_{2,n-1}}(L_{F_{n-1,k}})] + \frac{C}{n^5}, \end{aligned}$$

Fix an integer $k \geq \frac{n}{2} + 2\sqrt{n\ln n}$. We prove by induction on l that for all integer $l \leq \sqrt{n\ln n}$,

$$V_{\mu_{2,n}}(L_{F_{n,k}}) \leq \frac{1}{3^l}\sum_{i=0}^{l} C_l^i V_{\mu_{2,n-l}}(L_{F_{n-l,k-i}}) + C\sum_{i=0}^{l-1} \frac{1}{(n-i)^5}(\frac{2}{3})^i.$$

Indeed, if $l \leq \sqrt{n \ln n}$, then for all $i \in \{0, ..., l\}$,

$$k - i - \frac{n-l}{2} \geq \sqrt{n \ln n} \geq \sqrt{(n-l)\ln(n-l)},$$

therefore

$$V_{\mu_{2,n-l}}(L_{F_{n-l,k-i}}) \leq \frac{1}{3}(V_{\mu_{2,n-l-1}}(L_{F_{n-l-1,k-i-1}}) + V_{\mu_{2,n-l-1}}(L_{F_{n-l-1,k-i}})) + \frac{C}{(n-l)^5}.$$

Together with the induction hypothesis, this imply that

$$\begin{aligned}
V_{\mu_{2,n}}(L_{F_{n,k}}) &\leq \frac{1}{3^l}\sum_{i=0}^{l} C_l^i [\frac{1}{3}(V_{\mu_{2,n-l-1}}(L_{F_{n-l-1,k-i-1}}) + V_{\mu_{2,n-l-1}}(L_{F_{n-l-1,k-i}})) + \frac{C}{(n-l)^5}] \\
&+ C\sum_{i=0}^{l-1} \frac{1}{(n-i)^5}(\frac{2}{3})^i \\
&= \frac{1}{3^{l+1}}\{\frac{1}{3}(V_{\mu_{2,n-(l+1)}}(L_{F_{n-(l+1),k-(l+1)}}) + V_{\mu_{2,n-(l+1)}}(L_{F_{n-l-1,k}}) \\
&+ \sum_{j=1}^{l}(C_l^{j-1} + C_l^j)V_{\mu_{2,n-(l+1)}}(L_{F_{n-(l+1),k-j}})\} \\
&+ C\sum_{j=0}^{l} \frac{1}{(n-j)^5}(\frac{2}{3})^j \\
&= \frac{1}{3^{l+1}}\sum_{i=0}^{l+1} C_{l+1}^i V_{\mu_{2,n-(l+1)}}(L_{F_{n-(l+1),k-i}}) + C\sum_{j=0}^{l}\frac{1}{(n-j)^5}(\frac{2}{3})^j.
\end{aligned}$$

As before, it is easy to prove that

$$\sum_{j=0}^{l-1} \frac{1}{(n-j)^5}(\frac{2}{3})^j \leq \frac{C}{n^5}.$$

Furthermore (remember that $0 \leq L_F \leq 1$),

$$\frac{1}{3^l}\sum_{i=0}^{l} C_l^i V_{\mu_{2,n-l}}(L_{F_{n-l,k-i}}) \leq \frac{1}{3^l}\sum_{i=0}^{l} C_l^i = (\frac{2}{3})^l.$$

Thus, with $l = \sqrt{n \ln n}$ we find that

$$\begin{aligned}
V_{\mu_{2,n}}(L_{F_{n,k}}) &\leq \frac{1}{3^l}\sum_{i=0}^{l} C_l^i V_{\mu_{2,n-l}}(L_{F_{n-l,k-i}}) + C\sum_{i=0}^{l-1}\frac{1}{(n-i)^5}(\frac{2}{3})^i \\
&\leq (\frac{2}{3})^{\sqrt{n \ln n}} + \frac{C}{n^5} \\
&\leq \frac{C}{n^5}. \square
\end{aligned}$$

6.2 End of proof of theorem 2

Let k be an integer between $\frac{n}{2} - 2\sqrt{n \ln n}$ and $\frac{n}{2} + 2\sqrt{n \ln n}$. By proposition 3 and Chebyshev inequality, for all positive numbers ε,

$$(1) \quad \mu_{2,n}(\{p \in \Delta_n : \left|L_{F_{n,k}}(p) - C_n^k 2^{-n}\right| \geq \varepsilon\}) \leq \frac{C}{\varepsilon^2 n^2},$$

thus

$$\mu_{2,n}(\{p\in\Delta_n : \max_{\frac{n}{2}-2\sqrt{n\ln n}\le k\le \frac{n}{2}+2\sqrt{n\ln n}} \left|L_{F_{n,k}}(p)-C_n^k 2^{-n}\right|\ge\varepsilon\}) \le 4\sqrt{n\ln n}\frac{C}{\varepsilon^2 n^2}.$$

Let k be an integer in $\{0,...,n\}$ such that $\left|\frac{n}{2}-k\right|\ge 2\sqrt{n\ln n}$. By proposition 4 and Chebyshev inequality, for all positive numbers ε,

$$(2)\quad \mu_{2,n}(\{p\in\Delta_n : \left|L_{F_{n,k}}(p)-C_n^k 2^{-n}\right|\ge\varepsilon\}) \le \frac{C}{\varepsilon^2 n^5},$$

thus

$$\mu_{2,n}(\{p\in\Delta_n : \max_{k:\left|\frac{n}{2}-2\right|\ge 2\sqrt{n\ln n}} \left|L_{F_{n,k}}(p)-C_n^k 2^{-n}\right|\ge\varepsilon\}) \le n\frac{C}{\varepsilon^2 n^5}.$$

It follows that

$$\mu_{2,n}(\{p\in\Delta_n : \max_{k\in\{0,...,n\}} |L_n(p)(k)-L_n(\mathcal{B}_n)(p)|\le\varepsilon\}) \ge 1-\frac{C\ln n}{\varepsilon^2 n^{3/2}}.$$

Let δ be a positive number. We shall use inequality (1) with $\varepsilon_1=\frac{\delta}{4\sqrt{n\ln n}}$ and inequality (2) with $\varepsilon_2=\frac{\delta}{n}$. For each subset I of $\{0,...,n\}$, the set

$$\{p\in\Delta_n : |L_n(p)(I)-L_n(\mathcal{B}_n)(I)|\ge\delta\}$$

is included in the union of

$$\bigcup_{k:\left|\frac{n}{2}-k\right|\le 2\sqrt{n\ln n}} \{p\in\Delta_n : |L_n(p)(k)-L_n(\mathcal{B}_n)(k)|\ge\frac{\delta}{4\sqrt{n\ln n}}\}$$

and of

$$\bigcup_{k:\left|\frac{n}{2}-k\right|\ge 2\sqrt{n\ln n}} \{p\in\Delta_n : |L_n(p)(k)-L_n(\mathcal{B}_n)(k)|\ge\frac{\delta}{n}\}$$

which does not depend on I, therefore (remember that $L_n(p)(k)=L_{F_{n,k}}(p)$)

$$\mu_{2,n}(\{p\in\Delta_n : \sup_{I\subset\{1,...,n\}} |L_n(p)(I)-L_n(\mathcal{B}_n)(I)|\ge\delta)\}$$

$$\le \sum_{k:\left|\frac{n}{2}-k\right|\le 2\sqrt{n\ln n}} \mu_{2,n}(\{p\in\Delta_n : |L_n(p)(k)-L_n(\mathcal{B}_n)(k)|\ge\frac{\delta}{4\sqrt{n\ln n}}\})$$

$$+\sum_{k:\left|\frac{n}{2}-k\right|\ge 2\sqrt{n\ln n}} \mu_{2,n}(\{p\in\Delta_n : |L_n(p)(k)-L_n(\mathcal{B}_n)(k)|\ge\frac{\delta}{n}\})$$

$$\le \frac{C\sqrt{n\ln n}}{\varepsilon_1^2 n^2}+\frac{C}{\varepsilon_2^2 n^4} = \frac{C}{\delta^2}(\frac{16\sqrt{n}\ln^{3/2} n}{n}+\frac{1}{n^2}\}$$

$$\le \frac{C\ln^{3/2} n}{\sqrt{n}\delta^2}. \square$$

References

[Da, Du]: D. Dacunha-Castelle, M. Duflo, Exercices de probabilités et statistiques 1. Problèmes à temps fixe, Masson, Collection Mathématiques appliquées pour la Maîtrise, Paris p. 64-65, 1982

[Fe]: W. Feller, an introduction to probability theory and its applications, vol 1, 3^{th} ed., Wiley, 1968.

[Le]: E. Lesigne, Pile ou Face, Une Introduction au Calcul des Probabilités, ellipse, 2001.

[T,T]: K. Takeuchi, A. Takemura, *On sum of* $0-1$ *random variables 1. Univariate case*, Ann. Inst. Statist. Math. 39 (1987), Part A, 85-102.