Ann. I. H. Poincaré – PR 38, 6 (2002) 991–1007

S0246-0203(02)01130-5/FLA

TUSNADY'S LEMMA, 24 YEARS LATER

Pascal MASSART

Laboratoire de mathématique, UMR C 8628, bâtiment 425, centre d'Orsay, Université Paris-Sud, 91405 Orsay cedex, France

Received 29 June 2001, revised 15 May 2002

Dedicated to Jean Bretagnolle

ABSTRACT. – The optimal coupling between a variable with the $\mathrm{Bin}(n, 1/2)$ distribution and a normal random variable lies at the heart of the proof of the KMT Theorem for the empirical distribution function. Tusnády's Lemma (published in 1977 in his dissertation and in Hungarian) provides an inequality with explicit absolute constants which says that for this coupling, the distance between the random variables remains bounded in probability. In the appendix of a joint work with Jean Bretagnolle (1989), we have proposed a proof of Tusnády's Lemma which though elementary is highly technical and considered as rather obscure, at least this is what we have understood from several conversations with motivated readers. The purpose of this paper is to provide an alternative proof which is still based on elementary computations but which we hope to be simpler and more illuminating. This new proof also leads to a slight improvement on the original result in terms of constants.

MSC: primary 60F17, 60F99; secondary 62G30

Keywords: Quantile transform; Normal approximation; Binomial distribution; Hungarian construction

RÉSUMÉ. – Le couplage optimal d'une variable binomiale $\mathrm{Bin}(n, 1/2)$ et d'une variable gaussienne est au coeur de la preuve du Théorème de Komlós, Major et Tusnády pour la fonction de répartition empirique. Le Lemme de Tusnády (publié en 1977 dans sa thèse et en Hongrois) fournit une inégalité comportant des constantes absolues explicites, exprimant que l'écart entre ces variables convenablement couplées reste borné en probabilité. En appendice d'un article écrit en collaboration avec Jean Bretagnolle (1989), nous avons proposé une preuve du Lemme de Tusnády qui pour être élémentaire n'en est pas moins très technique et considérée comme plutôt obscure, c'est du moins ce que nous avons compris des quelques conversations que nous avons eues avec des lecteurs motivés. Le but de cet article est de proposer une nouvelle preuve, fondée elle aussi sur des calculs élémentaires mais que nous espérons plus simple et plus limpide. Cette preuve possède également le mérite de conduire à une amélioration (modeste) du résultat original au niveau des constantes.

E-mail address: Pascal.Massart@math.u-psud.fr (P. Massart).

1. Introduction

One of the most striking results on the empirical distribution function is the strong approximation by a Brownian bridge at an optimal rate due to Komlós, Major and Tusnády (see [5]). This celebrated result is now referred to as the KMT Theorem and the dyadic coupling scheme that Komlós, Major and Tusnády have introduced is often called the "Hungarian construction". The KMT Theorem is a very powerful tool which has been used in several papers devoted to the asymptotic behavior of nonparametric estimators (see for instance [8]). Quite recently, Nussbaum's works on strong approximation of experiments in Le Cam's sense in a nonparametric context (see especially [7]) have also stressed on the importance of the Hungarian construction by itself. As explained in the book by Csörgő and Révész [2] a crucial argument in the Hungarian construction is the coupling of a symmetric Binomial random variable with a normally distributed random variable with the same mean and variance. The proof of this step in the original paper by Komlós, Major and Tusnády is only a sketch but can be detailed as shown by Mason and van Zwet in [6]. An alternative proof of the Hungarian construction is proposed in [2]. It relies on a coupling inequality due to Tusnády. Since Tusnády's proof was not easily accessible because it only appeared in his Thesis [9] which is written in Hungarian, we have presented an alternative complete proof of it in the Appendix of a joint paper with Jean Bretagnolle (see [1]). It turns out that this proof, though elementary is rather intricate as noted by Carter and Pollard [3] in their recent attempt to produce a result which has the same flavor but which is not strictly comparable to Tusnády's inequality since it provides asymptotically a better estimate but with less precise absolute constants. Our aim in this paper is to propose a new proof of Tusnády's inequality. This new proof is built in the same spirit as the one that we originally provided in our joint paper with Bretagnolle but the line is (at least we hope!) simpler and it also leads to a (slightly) better result concerning the absolute constants. Before stating our result, let us recall that if Y has a continuous distribution function Φ and if F is a distribution function, denoting by F^{-1} the generalized inverse of the monotone function F, the *quantile transform* $F^{-1} \circ \Phi$ allows to define from Y a random variable $X = F^{-1} \circ \Phi(Y)$ with distribution function F. Our main result can be stated as follows.

THEOREM 1.1. – *Let Y be some standard normal random variable, n be some positive integer and B_n be the random variable with the symmetric Binomial distribution* $\mathrm{Bin}(n, 1/2)$, *defined from Y via the quantile transform. Then the following inequality holds*

$$\left| B_n - \frac{n}{2} - \frac{\sqrt{n}}{2} Y \right| \leqslant \frac{3}{4} + \frac{Y^2}{8}. \tag{1.1}$$

Note that the constant $3/4$ in the right hand side of (1.1) improves on the constant 1 appearing in Tusnády's original inequality. The proof of Theorem 1.1 easily derives from the following Gaussian comparisons for the Binomial tails. First we state an upper bound.

LEMMA 1.2. – *Let n be any positive integer. Let B_n and Y be respectively some* $\mathrm{Bin}(n, 1/2)$ *and standard normal random variables. Then, for every integer j such that*

$0 \leqslant j \leqslant n$ and $n+j$ is even, one has

$$\mathbb{P}\left[B_n \geqslant \frac{(n+j)}{2}\right] \leqslant \mathbb{P}\left[\frac{\sqrt{n}}{2}Y \geqslant \frac{j}{2} - \frac{3}{4}\right]. \tag{1.2}$$

Secondly, here is a lower bound.

LEMMA 1.3. – *Let B_n and Y be as in Lemma* 1.2 *above. For every integer j such that $0 \leqslant j \leqslant n$ and $n+j$ is even, the following inequality holds*

$$\mathbb{P}\left[B_n \geqslant \frac{(n+j)}{2}\right] \geqslant \mathbb{P}\left[\frac{\sqrt{n}}{2}Y \geqslant n\left(1 - \sqrt{1 - \frac{j}{n} + \frac{1}{2n}}\right)\right]. \tag{1.3}$$

Note that the gain with respect to Tusnády's original constant 1 (which becomes $3/4$ in our statement) will be obtained not only because $3/4$ appears in (1.2) instead of 1 but also with the help of the extra term $1/2n$ in (1.3). Our main task will be to prove these bounds for the Binomial tails. Let us see right now how they imply Theorem 1.1.

Proof of Theorem 1.1. – We denote by Φ, the distribution function of the standard normal distribution and consider some nonnegative integer j such that $n+j$ is even and $j \leqslant n$. We derive from (1.2) that

$$\mathbb{P}\left[B_n < \frac{(n+j)}{2}\right] \geqslant \Phi\left(\left(j - \frac{3}{2}\right)\Big/\sqrt{n}\right). \tag{1.4}$$

But using the well known inequality $\sqrt{1+u} \leqslant 1 + u/2$, $u > 0$, we have

$$2\left(-1 + \sqrt{1 + \frac{j}{n} - \frac{3}{2n}}\right) \leqslant \frac{j}{n} - \frac{3}{2n}$$

and therefore for any $j \in \mathbb{N}$ such that $n+j$ is even and $0 \leqslant j \leqslant n$, the following inequality holds

$$\mathbb{P}\left[B_n < \frac{(n+j)}{2}\right] \geqslant \Phi\left(\left(-2\sqrt{n} + 2\sqrt{n + j - \frac{3}{2}}\right)\right). \tag{1.5}$$

On the other hand, one derives by symmetry from (1.3) that for any integer j such that $n+j$ is even and $0 < j < n$ (note that j becomes $j-2$ when using 1.3)

$$\mathbb{P}\left[B_n < \frac{(n-j)}{2}\right] = \mathbb{P}\left[B_n > \frac{(n+j)}{2}\right] \geqslant \Phi\left(\left(-2\sqrt{n} + 2\sqrt{n - j - \frac{3}{2}}\right)\right). \tag{1.6}$$

Now (1.5) and (1.6) imply (since the d.f. of B_n is piecewise constant) that, for every $t \in \mathbb{R}$ such that $n + 2t - 3/2 \geqslant 0$, one has

$$\mathbb{P}\left[B_n \leqslant \frac{n}{2} + t\right] \geqslant \Phi\left(\left(-2\sqrt{n} + 2\sqrt{n + 2t - \frac{3}{2}}\right)\right)$$

or equivalently, for any $y \in \mathbb{R}$

$$\Phi(y) \leqslant \mathbb{P}\left[B_n \leqslant \frac{n}{2} + \frac{\sqrt{n}}{2} y + \frac{3}{4} + \frac{y^2}{8}\right].$$

This clearly implies by definition of the quantile transform that

$$B_n - \frac{n}{2} \leqslant \frac{\sqrt{n}}{2} Y + \frac{3}{4} + \frac{Y^2}{8},$$

which leads to Theorem 1.1 by symmetry. □

One could wonder whether the constant $3/4$ appearing in (1.2) can be improved. Obviously if (1.2) holds for some constant C instead of $3/4$, then C must be not smaller than $1/2$ (just look at the case $j = 1$ when n is odd). Inspecting a table of the standard normal distribution, it is also clear that $C = 1/2$ does not work but also that the "truth" is closer to $1/2$ than to $3/4$. Hence there is still some room to improve on our result. In our opinion this should be done in the spirit of Carter and Pollard [3] by taking C not as an absolute constant but rather as $C = (1/2) + (\theta/\sqrt{n})$ and find some adequate value for θ. This could be obtained by refining our technics but since we do not see how to get the corresponding improvement for the lower bound (1.3) we have decided not to present it here by sake of simplicity. Our approach for proving Lemma 1.2 and Lemma 1.3 will consist in summing up local comparisons between Binomial and Gaussian probabilities. The intuition coming from the usual Gaussian approximation of the Binomial with a correction of continuity, should be to compare the Binomial probability $p_n(k) = \mathbb{P}[2B_n = n + k]$ with $\mathbb{P}[k - 1 \leqslant \sqrt{n}Y \leqslant k + 1]$. The most delicate part of the game that we shall play below will be to design proper intervals $I_n(k)$'s on which it is relevant to compute the Gaussian probability $\mathbb{P}[\sqrt{n}Y \in I_n(k)]$ in order to get an easy comparison with $\mathbb{P}[2B_n = n + k]$. The definition of $I_n(k)$ will change according to wether we aim at getting a lower or an upper bound for $p_n(k)$. The easiest transformation that we shall use is a shift with length $1/2$ from the "intuitive choice" $[k - 1, k + 1]$ but we shall also use slightly more sophisticated transformations. This is what we shall study in the following section.

2. Nonasymptotic local expansions for binomials

Let us fix some notations that we shall use all along the paper.

DEFINITION 2.1. – *For any integer k, let $p_n(k) = 0$ if $n + k$ is odd and*

$$p_n(k) = \binom{n}{(n+k)/2} 2^{-n},$$

if $n + k$ is even. Moreover, for any $t \in (-1, 1)$, one defines

$$h(t) = (1 + t)\log(1 + t) + (1 - t)\log(1 - t),$$

$$f_n(t) = \sqrt{\frac{n}{2\pi}} \frac{1}{\sqrt{1-t^2}} \exp\left[-nh(t)/2\right]$$

and

$$g_n(t) = \sqrt{\frac{n}{2\pi}} \frac{1}{\sqrt{1-t}} \exp\left[-2n\left(1-\sqrt{1-t}\,\right)^2\right].$$

The functions f_n and g_n that we have just introduced will play the role of densities. This is clear for g_n for which one has for every $x \in [0,1]$

$$\mathbb{P}\left[0 \leqslant \frac{\sqrt{n}}{2} Y \leqslant n\left(1-\sqrt{1-x}\,\right)\right] = \int_0^x g_n(t)\,\mathrm{d}t, \tag{2.1}$$

where we recall that Y denotes a standard normal random variable. The relationship between f_n and Binomial probabilities comes from Stirling's formula. Indeed, let us first recall the classical nonasymptotic inequalities associated with Stirling's formula

$$1 + \frac{1}{12m} \leqslant \frac{m!}{(m/\mathrm{e})^m \sqrt{2\pi m}} \leqslant \exp\left(\frac{1}{12m}\right), \quad m \geqslant 1.$$

Then, using these inequalities, one easily derives that for any integer $k \leqslant n-2$ such that $n+k$ is even, one can write

$$p_n(k) = \frac{2}{n} f_n(k/n) C_n(k/n), \tag{2.2}$$

where the correction factor C_n due to the use of Stirling's formula obeys the following inequalities

$$\log C_n(x) \geqslant \log\left(1 + \frac{1}{12n}\right) - \frac{1}{3n(1-x^2)}$$

and

$$\log C_n(x) \leqslant \frac{1}{12n} - \log\left(1 + \frac{1}{6n(1-x)}\right) - \log\left(1 + \frac{1}{6n(1+x)}\right)$$

for any $x \in (0,1)$. Now, taking into account the monotonicity of $t \to (1/t)\log(1+t)$, setting $\gamma = 12 \log(13/12)$, one has for any $t \in (0, 1/12)$, $\log(1+t) \geqslant \gamma t$. Hence the above inequalities on the correction C_n become

$$\frac{\gamma}{12n} - \frac{1}{3n(1-x^2)} \leqslant \log C_n(x) \leqslant \frac{1}{12n} - \frac{\gamma}{3n(1-x^2)}, \quad \forall x \in \left[0, 1-\frac{2}{n}\right], \tag{2.3}$$

where the absolute constant γ satisfies $\gamma > 0.96051$. In view of (2.2), it is tempting to compare $p_n(k)$ with the integral of f_n on the interval $[(k-1)/n, (k+1)/n]$. This is essentially what is done below.

LEMMA 2.1. – *Let $n \geqslant 2$ and p_n, f_n be as in Definition* 2.1. *For any integer $k \leqslant n-2$ such that $n+k$ is even, the following inequalities are valid*

$$p_n(k) \leqslant \int_{(k-1)/n}^{(k+1)/n} f_n(t) \exp\left(\frac{1}{12n} - \frac{0.1479}{n(1-t^2)}\right) \mathrm{d}t. \tag{2.4}$$

If one moreover assumes that $k^2(1-k^2/n^2) \geqslant 2n/3$, one also has

$$p_n(k) \geqslant \int_{(k-1)/n}^{(k+1)/n} \exp\left(\frac{\gamma}{12n}\right) \sqrt{1-t^2}\, f_n(t)\, \mathrm{d}t, \tag{2.5}$$

where $\gamma = 12\log(13/12) > 0.96$.

Proof. – We use Lemma A.1 that we shall prove in Appendix A below. The proof of (2.5) is immediate from (2.2), (2.3) and (A.2). In order to prove (2.4), we notice that, given $\theta \in [0, 1/2]$ the logarithm ϕ of the function

$$t \to \frac{1}{\sqrt{1-t^2}} \exp\left(-\frac{\theta}{n(1-t^2)}\right)$$

is convex on $[-1+1/n, 1-1/n]$. Then, we apply Jensen's inequality (twice!), using first the convexity of the exponential and then the convexity of ϕ. Hence, setting $x = k/n$ and $\delta = 1/n$, a lower bound for

$$\frac{n}{2}\int_{(k-1)/n}^{(k+1)/n} f_n(t) \exp\left(-\frac{\theta}{n(1-t^2)}\right) \mathrm{d}t = \sqrt{\frac{n}{2\pi}} \times \frac{1}{2\delta} \int_{x-\delta}^{x+\delta} \exp\left(\phi(t) - \frac{n}{2}h(t)\right) \mathrm{d}t$$

is

$$\sqrt{\frac{n}{2\pi}} \exp\left(\phi(x) - \frac{n}{2} \times \frac{1}{2\delta} \int_{x-\delta}^{x+\delta} h(t)\, \mathrm{d}t\right),$$

which via (A.1) is bounded from below by

$$\sqrt{\frac{n}{2\pi}} \exp\left(\phi(x) - \frac{n}{2}h(x) - \frac{31}{180n(1-x^2)}\right)$$

and therefore

$$\frac{n}{2}\int_{(k-1)/n}^{(k+1)/n} f_n(t) \exp\left(-\frac{\theta}{n(1-t^2)}\right) \mathrm{d}t \geqslant f_n(x) \exp\left[-\frac{1}{n(1-x^2)}\left(\theta + \frac{31}{180}\right)\right].$$

Combining this lower bound with (2.2) and (2.3), yields by noticing that $\gamma/3 > 0.96051/3 > \theta + (31/180)$ holds true when choosing $\theta = 0.1479$

$$\int_{(k-1)/n}^{(k+1)/n} f_n(t)\exp\left(-\frac{\theta}{n(1-t^2)}\right)\mathrm{d}t \geqslant p_n(k)\exp\left(-\frac{1}{12n}\right).$$

Hence (2.4) is proven. □

We are in position to prove the Gaussian estimates on the Binomial probabilities which will lead to the Gaussian comparisons for the Binomial tails stated in Section 1. Now that the link between the binomial probabilities and f_n has been established and since (2.1) provides the link between g_n and the Gaussian probability involved in Lemma 1.3, we would like to provide a guideline for the proof of Lemma 1.3 which is the most delicate result to be proven below. Proving Lemma 1.3 essentially amounts (neglecting at this stage some remainder terms) to prove that for every $x \in [0, 1]$

$$\int_0^x f_n(t)\,\mathrm{d}t \leqslant \int_0^x g_n(t)\,\mathrm{d}t. \tag{2.6}$$

This inequality holds true for $x = 0$ and one can hope that it holds true for $x = 1$ because $\int_0^x f_n(t)\,\mathrm{d}t \simeq \int_0^x g_n(t)\,\mathrm{d}t \simeq 1/2$. Hence, by monotonicity, (2.6) will hold true if one is able to show that for some point t_n, $f_n(t) \leqslant g_n(t)$ whenever $t \leqslant t_n$ and $f_n(t) > g_n(t)$ otherwise. This is by essence the meaning of Proposition 2.4 below and we shall use a discrete version of the above argument to derive Lemma 1.3 from Proposition 2.4. This also explains why somehow surprisingly at first glance, a local upper bound for binomial probabilities such as Lemma 2.2 below will turn to be useful not only to prove the upper bound on the Binomial tail (1.2) but also to prove the lower bound (1.3) as well.

LEMMA 2.2. – *Let $n \geqslant 2$ and p_n be as in Definition* 2.1. *The following inequalities hold for every integer k such that $n + k$ is even and $k \leqslant n - 2$*

$$p_n(k) \leqslant \sqrt{\frac{n}{2\pi}} \int_{(k-1)/n}^{(k+1)/n} \exp\left(-\frac{(n-1)t^2}{2}\right)\mathrm{d}t. \tag{2.7}$$

Proof. – In order to prove (2.7), we use Lemma A.2 (see Appendix A) and more precisely (A.4) which ensures that for every $t \in (-1, 1)$ one has

$$-\frac{1}{2}\left[nh(t) + \log\left(1 - t^2\right) - (n-1)t^2\right] \leqslant \frac{t^{2n}}{4n(1-t^2)}.$$

Now, for $|t| \leqslant 1 - 1/n$, we have $t^{2n} \leqslant \mathrm{e}^{-2}$ which implies (2.7) via (2.4) and (A.4) since $\mathrm{e}^{-2}/4 \leqslant 0.1479 - (1/12)$. □

As a first consequence of this result, we derive the local estimate for the Binomial probabilities which will imply Lemma 1.2. Note that the calibration of the shift in the Gaussian probabilities below directly derives from (2.9).

PROPOSITION 2.3. – *Let $n \geqslant 4$, p_n be as in Definition* 2.1 *and Y be a standard normal random variable. Then, for any integer k such that $n+k$ is even and $2 \leqslant k \leqslant n-2$*

$$p_n(k) \leqslant \mathbb{P}\left[k - \frac{3}{2} \leqslant \sqrt{n}Y \leqslant \left(k + \frac{1}{2}\right)\right]. \tag{2.8}$$

Proof. – In order to bound the binomial probabilities $p_n(k)$, for $2 \leqslant k \leqslant n-2$, we use (2.7). We note that since $t \to -t^2 + t - (1/4n)$ is concave and positive at points $t = 1/n$ and $t = 1-(1/n)$, it is positive on the whole interval $[1/n, 1-(1/n)]$, which means that for every t belonging to this interval one has

$$(n-1)t^2 \geqslant n\left(t - \frac{1}{2n}\right)^2. \tag{2.9}$$

Hence we derive from (2.7) that for $2 \leqslant k \leqslant n-2$ (and $n+k$ even) an upper bound for $\sqrt{2\pi/n}\; p_n(k)$ is given by

$$\int_{(k-1)/n}^{(k+1)/n} \exp\left(-n\left(t - \frac{1}{2n}\right)^2 / 2\right) \mathrm{d}t \leqslant \int_{(k-(3/2))/n}^{(k+(1/2))/n} \exp\left(-nu^2/2\right) \mathrm{d}u$$

and the result follows. □

We turn now to the local comparisons from which we shall derive Lemma 1.3.

PROPOSITION 2.4. – *Let $n \geqslant 4$ and p_n, g_n be as in Definition* 2.1

$$p_n(k) \leqslant \int_{(k-1)/n}^{(k+1)/n} g_n(t)\,\mathrm{d}t, \quad \textit{if } \frac{1}{n} \leqslant \frac{k}{n} \leqslant \frac{1}{\sqrt{n}} - \frac{1}{n}, \tag{2.10}$$

$$p_n(k) \leqslant 2\int_0^{1/n} g_n(t)\,\mathrm{d}t, \quad \textit{if } k = 0 \tag{2.11}$$

and if $n \geqslant 5$

$$p_n(k) \geqslant \int_{(k-(1/2))/n}^{(k+(3/2))/n} g_n(t)\,\mathrm{d}t, \quad \textit{whenever } \frac{1}{\sqrt{n}} + \frac{1}{2n} \leqslant \frac{k}{n} \leqslant 1 - \frac{2}{n}. \tag{2.12}$$

Proof. – In view of (2.7), in order to prove (2.10) and (2.11), it is enough to prove that for every nonnegative $t \leqslant 1/\sqrt{n}$ one has

$$\exp\left(-\frac{(n-1)t^2}{2}\right) \leqslant \frac{1}{\sqrt{1-t}} \exp\left(-2n\left(1 - \sqrt{1-t}\right)^2\right).$$

By (A.5) this will a fortiori hold if

$$\frac{t^2}{2} - \frac{nt^2}{2} \leqslant -\frac{1}{2}\log(1-t) - nt^2\left(\frac{1}{2} - \frac{t}{4} - \frac{1}{2}\log(1-t)\right)$$

or equivalently

$$\frac{t^2}{2} \leqslant -\frac{1}{2}\log(1-t)\left(1-nt^2\right) + \frac{nt^3}{4}. \tag{2.13}$$

Now $nt^2 \leqslant 1$ and $-\log(1-t) \geqslant t$ hence

$$-\frac{1}{2}\log(1-t)\left(1-nt^2\right) + \frac{nt^3}{4} \geqslant \frac{t}{2} - \frac{nt^3}{4} - \frac{t^2}{2} \geqslant \frac{t}{4} - \frac{t^2}{2}$$

which is nonnegative because $t \leqslant 1/\sqrt{n} \leqslant 1/2$. This shows that (2.13) is true and therefore (2.10) and (2.11) hold.

Let us turn to the proof of (2.12). We derive from (2.5) that

$$p_n(k) \geqslant \sqrt{\frac{n}{2\pi}} \int_{(k-(1/2))/n}^{(k+(3/2))/n} \exp\left(\frac{\gamma}{12n}\right) \exp\left(-\frac{nh(t-\frac{1}{2n})}{2}\right) \mathrm{d}t.$$

Now since the second derivative of h is nondecreasing on $(0,1)$, Taylor's formula implies that

$$\begin{aligned} h\left(t-\frac{1}{2n}\right) - h(t) &\leqslant -\frac{1}{2n}h'(t) + \frac{1}{8n^2}h''(t) \\ &\leqslant -\frac{1}{2n}\left[\log(1+t) - \log(1-t)\right] + \frac{1}{4n^2(1-t^2)} \end{aligned}$$

and therefore a lower bound for $p_n(k)$ is given by

$$\sqrt{\frac{n}{2\pi}} \int_{(k-(1/2))/n}^{(k+(3/2))/n} \exp\left[\frac{\gamma}{12n} + \frac{1}{4}\left(\log(1+t) - \log(1-t)\right) - \frac{n}{2}h(t)\right] \mathrm{d}t.$$

Hence, (2.12) will hold if we can check that the following inequality holds true for every t such that $1/\sqrt{n} \leqslant t \leqslant 1-(1/2n)$

$$\frac{1}{4}\log\left(1-t^2\right) + \frac{\gamma}{12n} - \frac{1}{8n(1-t^2)} + n\left[2\left(1-\sqrt{1-t}\,\right)^2 - \frac{h(t)}{2}\right] \geqslant 0.$$

Using (A.6) which is proved in Appendix A (and multiplying by 4), we see that the above inequality derives from

$$A(n,t) = \log\left(1-t^2\right) + \frac{\gamma}{3n} - \frac{1}{2n(1-t^2)} + nt^3 + \frac{7nt^4}{24} \geqslant 0 \tag{2.14}$$

and therefore it remains to show that (2.14) holds. Expanding $-\log(1-t^2)$ in power series one easily gets

$$-\log\left(1-t^2\right) \leqslant \frac{t^2}{2} + \frac{t^2}{2(1-t^2)},$$

which yields

$$A(n,t) \geqslant -\frac{t^2}{2} - \frac{t^2}{2(1-t^2)} + \frac{\gamma}{3n} - \frac{1}{2n(1-t^2)} + nt^3 + \frac{7nt^4}{24}.$$

But $-t^2(1/2) + (\gamma/3n) + t^4(7n/24) \geqslant 0$ for all t (because the discriminant is negative), hence

$$A(n,t) \geqslant -\frac{t^2}{2(1-t^2)} - \frac{1}{2n(1-t^2)} + nt^3$$

and since $nt^2(1-t^2) \geqslant 3/4$ (just use $1/\sqrt{n} \leqslant t \leqslant 1-(1/2n)$ and $n \geqslant 5$), one also has

$$\left(1-t^2\right)A(n,t) \geqslant \frac{3t}{4} - \frac{t^2}{4} - \frac{1}{2n}.$$

To check the positivity of the right hand side of this inequality, we notice that it is concave and that its values at the end points $t = 1/\sqrt{n}$ and $t = 1-(1/2n)$ are easily seen to be positive because $n \geqslant 5$. Hence $A(n,t)$ is positive for every t such that $1/\sqrt{n} \leqslant t \leqslant 1-(1/2n)$ and the result follows. □

We now pass from the above evaluations of the Binomial probabilities to Gaussian comparisons and establish the main results of this paper.

3. Upper bound for the Binomial tail

We can first easily derive from Proposition 2.3, the comparison between the Binomial tail and the corresponding Gaussian at a point which is shifted from a quantity which is slightly greater than the correction of continuity.

Proof of Lemma 1.2. – The result being trivial for $n=1$, we suppose that $n \geqslant 2$. We shall check (1.2) for the extremal values of j and then, if $n \geqslant 4$, bound each Binomial probability $\mathbb{P}[2B_n = n+k] = p_n(k)$ by the corresponding Gaussian probability that $\sqrt{n}Y$ belongs to the interval $[k-(3/2), k+(1/2)]$ for every k such that $2 \leqslant k \leqslant n-2$. Note here that these intervals do not overlap when k varies because of the constraint: $n+k$ is even. In order to check (1.2) when $j=n$, we use the classical lower bound for the standard Gaussian tail (see for instance [4], p. 17)

$$\mathbb{P}[Y \geqslant \lambda] \geqslant \frac{1}{\sqrt{2\pi}} \times \frac{1}{\sqrt{\lambda^2+2}} \exp\left(-\lambda^2/2\right)$$

at point $\lambda = \sqrt{n}(1-(3/2n))$. Noticing that $\lambda^2 \geqslant n-2$, this gives

$$\mathbb{P}\left[Y \geqslant \sqrt{n}\left(1-\frac{3}{2n}\right)\right] \geqslant \frac{1}{\sqrt{2\pi n}} \exp\left(-\frac{n}{2}+1\right) \geqslant \frac{1}{\sqrt{n}} \exp\left(-\frac{n}{2}\right)$$

which easily leads to

$$\mathbb{P}\left[Y \geqslant \sqrt{n}\left(1-\frac{3}{2n}\right)\right] \geqslant 2^{-n}$$

and therefore (1.2) when $j=n$. Assuming first that $n \geqslant 4$ and $2 \leqslant j \leqslant n-2$, (1.2) follows by summing up (2.8) "from the tail" that is for those indices k lying between j and $n-2$, using that (1.2) holds true at the end point $j=n$. It remains now to deal with the smallest value of j. When n is odd and $j=1$, (1.2) is trivial (since the right hand side is larger than $1/2$). When n is even and $j=0$, checking (1.2) amounts to verify that

$$p_n(0)/2 \leqslant \mathbb{P}\left[0 \leqslant Y \leqslant \frac{3}{2\sqrt{n}}\right]. \tag{3.1}$$

Now on the one hand, it comes from (2.2) and (2.3) that

$$p_n(0)/2 \leqslant \frac{1}{\sqrt{2\pi n}} \exp\left(-\frac{(4\gamma-1)}{12n}\right)$$

and on the other hand by Jensen's inequality

$$\int_0^\delta \exp\left(-\frac{x^2}{2}\right) \mathrm{d}x \geqslant \delta \exp\left(-\frac{1}{\delta}\int_0^\delta \frac{x^2}{2}\,\mathrm{d}x\right) = \delta \exp\left(-\frac{\delta^2}{6}\right),$$

which leads for $\delta = (3/2\sqrt{n})$ to

$$\mathbb{P}\left[0 \leqslant Y \leqslant \frac{3}{2\sqrt{n}}\right] \geqslant \frac{3}{2\sqrt{n}} \times \frac{1}{\sqrt{2\pi}} \exp\left(-\frac{3}{8n}\right).$$

Hence,

$$2\mathbb{P}\left[0 \leqslant Y \leqslant \frac{3}{2\sqrt{n}}\right] \geqslant p_n(0)\left[\frac{3}{2}\exp\left(-\frac{1}{n}\left(-\frac{1}{12}-\frac{3}{8}+\frac{\gamma}{3}\right)\right)\right] \geqslant p_n(0)$$

which means that (3.1) holds. This completes the proof when $n \geqslant 4$. For $n \leqslant 3$, there is nothing more to do than checking the inequality at the two extreme values for j and therefore (1.2) also holds in this case. □

4. Lower bound for the Binomial tail

The proof of Lemma 1.3 is a bit more delicate than that of Lemma 1.2 but is still quite easily obtainable from our local nonasymptotic expansions for the Binomial probabilities.

Proof of Lemma 1.3. – Assume first that (1.3) holds true for the extreme values $j=n$ and $j=0$ or $j=1$ according to the parity of n. Then, for the other values of j, (1.3) follows from Proposition 2.4 either by summing up (2.12) "from the tail", i.e. over indices k such that $j \leqslant k \leqslant n-2$ if $\frac{1}{\sqrt{n}} + \frac{1}{2n} \leqslant \frac{j}{n}$ and then use the identity

$$\begin{aligned}
&\mathbb{P}\left[\frac{\sqrt{n}}{2}Y \geqslant n\left(1-\sqrt{1-\frac{j}{n}+\frac{1}{2n}}\right)\right] \\
&\quad = \mathbb{P}\left[Y \geqslant 2\sqrt{n}\left(1-\sqrt{1/2n}\right)\right] + \int\limits_{(j-(1/2))/n}^{1-(1/2n)} g_n(t)\,\mathrm{d}t
\end{aligned}$$

or else by summing up (2.10) (using also (2.11) if n is even) “from the beginning”, i.e. over indices $k \leqslant j$ and then use the identities

$$\frac{p_n(0)}{2} + \sum_{1\leqslant k<j} p_n(k) = \frac{1}{2} - \mathbb{P}\left[B_n \geqslant \frac{(n+j)}{2}\right]$$

and by (2.1)

$$\int\limits_{0}^{(j-(1/2))/n} g_n(t)\,\mathrm{d}t = \frac{1}{2} - \mathbb{P}\left[\frac{\sqrt{n}}{2}Y \geqslant n\left(1-\sqrt{1-\frac{j}{n}+\frac{1}{2n}}\right)\right].$$

It remains therefore to check (1.3) at the end points.We begin by checking that (1.3) holds at point $j = n$. To do this we use the classical upper bound for the standard Gaussian tail

$$\mathbb{P}[Y \geqslant \lambda] \leqslant \frac{1}{2}\exp\left(-\frac{\lambda^2}{2}\right)$$

for $\lambda = 2\sqrt{n}(1-\sqrt{1/2n})$. This gives

$$\log \mathbb{P}[Y \geqslant \lambda] \leqslant -\log(2) - 2n + 2\sqrt{2n} - 1$$

and therefore

$$-n\log(2) - \log \mathbb{P}[Y \geqslant \lambda] \geqslant n\big(2-\log(2)\big) - 2\sqrt{2n} + 1 + \log(2).$$

But the discriminant of the polynomial $X^2(1-\log(2)/2) - 2X + 1 + \log(2)$ is negative, hence this polynomial is positive for every value of X and in particular for $X = \sqrt{2n}$ which implies the positivity of $-n\log(2) - \log \mathbb{P}[Y \geqslant \lambda]$. Hence $\mathbb{P}[Y \geqslant 2\sqrt{n}(1-\sqrt{1/2n})] \leqslant 2^{-n}$, which means that (1.3) holds for $j = n$. If $j = 0$, we have to prove that

$$p_n(0)/2 \geqslant \mathbb{P}\left[0 \leqslant Y \leqslant 2\sqrt{n}\left(\sqrt{1+\frac{1}{2n}}-1\right)\right]. \tag{4.1}$$

But

$$p_n(0)/2 \geqslant \exp\left(-\frac{1}{3n}\right)\frac{1}{\sqrt{2\pi n}} \geqslant \frac{1}{2\sqrt{2\pi n}} \geqslant \mathbb{P}\left[0 \leqslant Y \leqslant \frac{1}{2\sqrt{n}}\right],$$

which implies (4.1) since $\sqrt{1+1/2n}-1 \leqslant 1/4n$. The case $j = 1$ being trivial, the proof is now complete. $\square$

Acknowledgement

The author gratefully thanks an anonymous referee for his care and helpful suggestions.

Appendix A

We record here several elementary inequalities which have been used throughout the paper. Although each of these results can be checked quite easily by using adequate manipulations on Taylor expansions, we have chosen to present complete and detailed proofs to make the verifications easier to the motivated reader. Most of these results concern the function h of Definition 2.1. We begin with nonasymptotic bounds associated with the approximation of the meanvalue of a function on some interval by its value at the midpoint. These bounds are established first for the function h itself. Since it is convex on $(-1,1)$, the meanvalue of h on every subsegment $[x-\delta, x+\delta]$ of $(-1,1)$ is larger than $h(x)$ and the purpose of the following Lemma is to propose an upper bound on the error that one makes when approximating the mean value of h on $[x-\delta, x+\delta]$ by $h(x)$. In the same spirit, some corrective factor for the approximation of the mean value of $\exp(-h/2\delta)$ by $\exp(-h(x)/2\delta)$ is also provided, at least when x is large enough as compared to δ.

LEMMA A.1. – *For every positive number $\delta \leqslant 1/2$, one has for any x such that $0 \leqslant x \leqslant 1-2\delta$*

$$\frac{1}{2\delta}\int_{x-\delta}^{x+\delta} h(t)\,\mathrm{d}t \leqslant h(x) + \frac{31}{90}\frac{\delta^2}{1-x^2}. \tag{A.1}$$

If moreover $x^2(1-x^2) \geqslant 2\delta/3$, one also has

$$\frac{1}{2\delta}\int_{x-\delta}^{x+\delta} \exp\left[-\frac{h(t)}{2\delta}\right]\mathrm{d}t \leqslant \frac{1}{\sqrt{1-x^2}}\exp\left[-\frac{h(x)}{2\delta} - \frac{\delta}{3(1-x^2)}\right]. \tag{A.2}$$

Proof. – Expanding in power series, we get

$$\begin{aligned}\frac{1}{2\delta}\int_{x-\delta}^{x+\delta} h(t)\,\mathrm{d}t &= h(x) + \sum_{k=1}^{\infty}\frac{h^{(2k)}(x)}{(2k+1)!}\delta^{2k}\\ &= h(x) + \sum_{k=1}^{\infty}\frac{(2k-2)!}{(2k+1)!}\left[\frac{1}{(1-x)^{2k-1}} + \frac{1}{(1+x)^{2k-1}}\right]\delta^{2k}\\ &= h(x) + \frac{1}{3}\frac{\delta^2}{(1-x^2)} + R\end{aligned}$$

where the remainder term R is given by

$$R = \sum_{k=2}^{\infty}\frac{(2k-2)!}{(2k+1)!}\left[\frac{1}{(1-x)^{2k-1}} + \frac{1}{(1+x)^{2k-1}}\right]\delta^{2k}.$$

Now, on the one hand for any integer $k \geqslant 2$, $(2k-2)!/(2k+1)! \leqslant 2!/5! = 1/60$ and on the other hand, for any $\theta \in [0, 1/2]$ one has

$$\sum_{k=2}^{\infty} \theta^{2k-1} = \frac{\theta^3}{1-\theta^2} \leqslant \frac{\theta}{3}.$$

Hence, recalling that $\delta/(1-x) \leqslant 1/2$,

$$\begin{aligned} R &\leqslant \frac{\delta}{60} \sum_{k=2}^{\infty} \frac{\delta^{2k-1}}{(1-x)^{2k-1}} + \frac{\delta^{2k-1}}{(1+x)^{2k-1}} \leqslant \frac{\delta^2}{180}\left[\frac{1}{1-x} + \frac{1}{1+x}\right] \\ &\leqslant \frac{\delta^2}{90(1-x^2)} \end{aligned}$$

and (A.1) follows. We now turn to the proof of (A.2). We first notice that setting

$$I = \frac{1}{2\delta} \int_{x-\delta}^{x+\delta} \exp\left(-\frac{h(t)}{2\delta}\right) \mathrm{d}t,$$

one has

$$I = \frac{1}{2\delta} \int_{0}^{\delta} \left[\exp\left(-\frac{h(x+u)}{2\delta}\right) + \exp\left(-\frac{h(x-u)}{2\delta}\right)\right] \mathrm{d}u. \tag{A.3}$$

The Taylor expansion of h at point x can be written as

$$h(x+t) = h(x) + \sum_{j \geqslant 1} \frac{h^{(j)}(x)}{j!} t^j,$$

with $h'(x) = \log(1+x) - \log(1-x)$ and for any integer $j \geqslant 2$

$$h^{(j)}(x) = (j-2)!\left[(-1)^j (1+x)^{-j+1} + (1-x)^{-j+1}\right].$$

Of course the derivatives of the function h of any order j are nonnegative at every nonnegative point x. Hence when t is nonnegative, one has

$$h(x+t) \geqslant h(x) + th'(x) + \frac{t^2}{2} h''(x) \geqslant h(x) + th'(x) + \frac{t^2}{1-x^2}.$$

Getting a lower bound for $h(x+t)$ when $t = -u$ with $u \in [0, \delta]$ is slightly more complicated because we have to handle nonpositive terms in the Taylor expansion of h corresponding to the derivatives of odd orders. We therefore have to group consecutive even and odd terms. More precisely for any integer k, we note that $h^{(2k+1)}(x)/h^{(2k)}(x) \leqslant (2k-1)/(1-x)$ which leads to

$$\frac{h^{(2k)}(x)}{(2k)!} u^{2k} - \frac{h^{(2k+1)}(x)}{(2k+1)!} u^{2k+1} \geqslant \frac{h^{(2k)}(x)}{(2k)!} u^{2k}\left[1 - \frac{(2k-1)u}{(2k+1)(1-x)}\right]$$

which in turn implies since $u \leqslant (1-x)/2$

$$\frac{h^{(2k)}(x)}{(2k)!}u^{2k} - \frac{h^{(2k+1)}(x)}{(2k+1)!}u^{2k+1} \geqslant \frac{h^{(2k)}(x)}{(2k)!}u^{2k}\left[1 - \frac{(2k-1)}{2(2k+1)}\right] \geqslant 0.$$

Hence

$$h(x-u) \geqslant h(x) - uh'(x) + \frac{5u^2}{12}h''(x) \geqslant h(x) + th'(x) + \frac{5u^2}{6(1-x^2)}$$

and therefore whatever $|t| \leqslant \delta$, the following lower bound is a fortiori valid

$$h(x+t) \geqslant h(x) + th'(x) + \frac{2t^2}{3(1-x^2)}.$$

Plugging this lower bound in (A.3), we derive that

$$I \exp\left(\frac{h(x)}{2\delta}\right) \leqslant \frac{1}{\delta}\int_0^\delta \cosh\left(\frac{uh'(x)}{2\delta}\right)\exp\left(-\frac{u^2}{3\delta(1-x^2)}\right)\mathrm{d}u$$

which, since $\cosh(h'(x)/2) = (1-x^2)^{-1/2}$, will imply (A.2) if we can prove that the function

$$\psi : u \to \log\left[\cosh\left(\frac{uh'(x)}{2\delta}\right)\right] - \frac{u^2}{3\delta(1-x^2)}$$

is nondecreasing on $[0, \delta]$. To prove this we compute the derivative of ψ

$$\psi'(u) = \frac{h'(x)}{2\delta}\tanh\left(\frac{uh'(x)}{2\delta}\right) - \frac{2u}{3\delta(1-x^2)}.$$

Now, $\tanh(s)/s$ is nonincreasing on $\mathbb{R}_+$ so the nonnegativity of ψ' on $[0, \delta]$ has to be checked only at point δ. But $\tanh(h'(x)/2) = x$ and $h'(x)/2 \geqslant x$, so

$$\psi'(\delta) = \frac{xh'(x)}{2\delta} - \frac{2}{3(1-x^2)} \geqslant \frac{x^2}{\delta} - \frac{2}{3(1-x^2)}$$

which ensures that $\psi'(\delta)$ is nonnegative because of the assumption that $x^2(1-x^2) \geqslant 2\delta/3$. □

The following inequalities make easy the comparison of h with various quantities related to Gaussian log-probabilities.

LEMMA A.2. – *For every integer $n \geqslant 1$ and every $t \in (-1, 1)$, the following inequality holds*

$$-\frac{1}{2}\left[nh(t) + \log\left(1-t^2\right) - (n-1)t^2\right] \leqslant \frac{t^{2n}}{4n(1-t^2)}. \tag{A.4}$$

Moreover, for every $t \in (0,1)$, the following inequalities are valid

$$\frac{2}{t^2}\left(1-\sqrt{1-t}\right)^2 \leqslant \frac{1}{2}-\frac{t}{4}-\frac{1}{2}\log(1-t) \tag{A.5}$$

and

$$\frac{1}{t^2}\left[2\left(1-\sqrt{1-t}\right)^2-\frac{h(t)}{2}\right] \geqslant \frac{t}{4}+\frac{7t^2}{96}. \tag{A.6}$$

Proof. – Let us expand the left hand side $L(t)$ of (A.4) in power series. One has $L(t) = L_1(t) + L_2(t)$, where

$$L_1(t) = -\frac{1}{2}\left[\sum_{j=2}^{n-1}\frac{n}{j(2j-1)}t^{2j} - \sum_{j=2}^{n-1}\frac{1}{j}t^{2j}\right]$$

(if $n < 3$, one takes $L_1 = 0$) and

$$L_2(t) = -\frac{1}{2}\left[\sum_{j\geqslant n}\frac{n}{j(2j-1)}t^{2j} - \sum_{j\geqslant n}\frac{1}{j}t^{2j}\right].$$

Then, changing j into $n+1-j$ for those summation indices j which are larger than $(n+1)/2$, we get

$$L_1''(t) = -\sum_{j=2}^{n-1} t^{2j-2}(n+1-2j) = -\sum_{2\leqslant j\leqslant (n+1)/2}(n+1-2j)\left(t^{2j-2}-t^{2n-2j}\right) \leqslant 0$$

and therefore, since L_1 and its derivative are null at point 0, we conclude that L_1 is nonpositive. Finally

$$L_2(t) = \frac{1}{2}\left[\sum_{j\geqslant n}\frac{1}{j}\left(1-\frac{n}{(2j-1)}\right)t^{2j}\right] \leqslant \frac{1}{4n}\sum_{j\geqslant n}t^{2j}$$

and (A.4) follows. The power series expansion of the left hand side of (A.5) writes

$$\frac{1}{2}+\frac{t}{4}+\frac{5t^2}{32}+\sum_{j\geqslant 3}\frac{3\times\cdots\times(2j+1)}{(j+2)!2^j}t^j \tag{A.7}$$

as for the right hand side, one has

$$\frac{1}{2}+\frac{t}{4}+\frac{t^2}{4}+\sum_{j\geqslant 3}\frac{1}{2j}t^j. \tag{A.8}$$

For every integer $j \geqslant 3$, let us consider the ratio r_j between the term of degree j in (A.7) and the corresponding term in (A.8). Then

$$\frac{r_{j+1}}{r_j} = \frac{(2j+3)(j+1)}{2(j+3)j} = \frac{2j^2+5j+3}{2j^2+6j} \leqslant 1, \quad \text{for every } j \geqslant 3,$$

which means that the sequence $(r_j)_{j\geqslant 3}$ is nonincreasing and therefore $r_j \leqslant r_3 < 1$. Hence, each term of the expansion (A.7) is not larger than the term with the same degree in (A.8) which proves (A.5). To prove (A.6), we simply substract term by term to expansion (A.7) the power series expansion of $h(t)/(2t^2)$ which is

$$\frac{h(t)}{2t^2} = \frac{1}{2} + \frac{t^2}{12} + \sum_{k\geqslant 2} \frac{1}{(2k+2)(2k+1)} t^{2k}$$

and conclude since

$$\frac{3\times\cdots\times(2j+1)}{(j+2)!2^j} = \frac{3\times\cdots\times(2j+1)}{2\times\cdots\times 2j} \times \frac{1}{(j+1)(j+2)} \geqslant \frac{1}{(j+1)(j+2)}$$

for every even integer j larger than 3. $\square$

REFERENCES

[1] J. Bretagnolle, P. Massart, Hungarian constructions from the nonasymptotic view point, Ann. Probab. 17 (1) (1989) 239–256.

[2] M. Csörgő, P. Révész, Strong Approximations in Probability and Statistics, Academic Press, New York, 1981.

[3] A. Carter, D. Pollard, Tusnády's inequality revisited, Preprint, 2000.

[4] K. Ito, H.P. Mac Kean, Diffusion Processes and Their Sample Paths, Springer, Berlin, 1965.

[5] J. Komlós, P. Major, G. Tusnády, An approximation of partial sums of independent rv-s, and the sample df. I, Z. Warscheinlichkeitstheorie Verwandte Geb. 32 (1975) 111–131.

[6] D. Mason, W. van Zwet, A refinement of the KMT inequality for the uniform empirical process, Ann. Probab. 15 (1987) 871–884.

[7] M. Nussbaum, Asymptotic equivalence of density estimation and Gaussian white noise, Ann. Statist. 24 (1996) 2399–2430.

[8] G.R. Shorack, J.A. Wellner, Empirical Processes with Applications to Statistics, Wiley, New York, 1986.

[9] G. Tusnády, A study of statistical hypotheses, PhD Thesis, Hungarian Academy of Sciences, Budapest, 1977 (In Hungarian).