

Chapter 1: *tails of the normal distribution*

Pietro Maria Sparago and Shakeel Gavioli-Akilagun

Last updated: May 29 2024

1 Introduction

2 Tails of the normal distribution

Theorem 2.1. *Let $X \sim \mathcal{N}(0, 1)$. Define $\bar{\Phi}(x) := P(X > x), x \geq 0$. Then*

$$\frac{1}{2}e^{-x\sqrt{2/\pi}-x^2/2} \leq \bar{\Phi}(x) \leq \min\left(\frac{1}{2}e^{-x^2/2}, \frac{1}{x}\phi(x)\right), x \geq 0 \quad (1)$$

Proof. We have, by continuity and Corollary 2.1:

$$2\bar{\Phi}(x) = \lim_{y \downarrow 0} \frac{\bar{\Phi}(x+y)}{\bar{\Phi}(y)} \begin{cases} \leq \lim_{y \downarrow 0} e^{-xy-x^2/2} = e^{-x^2/2} \\ \geq \lim_{y \downarrow 0} e^{-\rho(y)x-x^2/2} = e^{-x\sqrt{2/\pi}-x^2/2} \end{cases} \quad (2)$$

We conclude for the upper bound with Lemma 2.1. □

Remark 2.1. *Note that Proposition 2.1 yields a stronger version of the infimum Chernoff bound which is found with*

$$\bar{\Phi}(x) \leq \inf_{\lambda \geq 0} e^{-\lambda x} e^{\lambda^2/2} = e^{-x^2/2}, x \geq 0$$

where the second equality follows from the quadratic minimization condition $\frac{d}{d\lambda}(-\lambda x + \lambda^2/2) = 0$ which yields $\lambda = x$.

2.1 Lemmas for the proof of Theorem 2.1 - mostly from Pollard (2002); Appendix D

Lemma 2.1. Let $\phi(x) := (2\pi)^{-1/2}e^{-x^2/2}$. Then:

$$\left(\frac{1}{x} - \frac{1}{x^3}\right)\phi(x) < \bar{\Phi}(x) < \frac{1}{x}\phi(x), \quad x > 0 \quad (3)$$

Proof. We have $\bar{\Phi}(x) = \int_x^\infty \phi(t)dt$. Note that for all $t > 0$ the functions $1 - 3/t^4 < 1$ and $1 + 1/t^2 > 1$ so $(1 - 3/t^4)\phi(t) < \phi(t) < (1 + 1/t^2)\phi(t)$. We thus integrate from $x > 0$ to infinity and obtain the claimed bounds. \square

It follows immediately that for the function $\rho(x) := \phi(x)/\bar{\Phi}(x)$, $x > 0$ we have the bounds

$$x < \rho(x) < \frac{x^3}{x^2 - 1}, \quad x > 1$$

Theorem 2.2. ρ is increasing, $\rho(-\infty) = 0$ and $\rho(0) = \sqrt{2/\pi}$. The function $\rho(x) - x$ decreases to zero as x tends to infinity. The function $\ln \rho(x)$ is concave and $\ln \rho(x + \delta) < \ln \rho(x) + (\rho(x) - x)\delta$ for $x \in \mathbb{R}$ and $\delta > 0$.

Proof. We have

$$\begin{aligned} \frac{1}{\rho(x)} &= \frac{\bar{\Phi}(x)}{\phi(x)} = \int_x^\infty \frac{\phi(t)}{\phi(x)} dt \\ &\stackrel{z=t-x}{=} \int_0^\infty \frac{\phi(z+x)}{\phi(x)} dz \\ &= \int_0^\infty e^{-zx} e^{-z^2/2} dz \end{aligned}$$

That is, the Laplace transform of the finite measure $\mu(dz) := e^{-z^2/2}dz, z > 0$. From this it follows that $\rho(0) = \sqrt{2/\pi}$ and that $\rho(-\infty) = 0$, since by monotone convergence $\lim_{x \rightarrow -\infty} \frac{1}{\rho(x)} = \infty$. For $x \geq 0$ we have $ze^{-zx} \leq z, z > 0$ and $z \in L^1(d\mu)$. So by differentiation-under-integral lemma we have

$$\begin{aligned} \rho(x) \left(\frac{d}{dx} \frac{1}{\rho(x)} \right) &= \rho(x) \left(- \int_0^\infty ze^{-zx} \mu(dz) \right) < 0 \\ \rho(x) \left(\frac{d^2}{dx^2} \frac{1}{\rho(x)} \right) &= \rho(x) \int_0^\infty z^2 e^{-zx} \mu(dz) > 0 \end{aligned}$$

The function $1/\rho$ is decreasing because $\frac{d}{dx} \frac{1}{\rho(x)} < 0$, so that ρ is increasing. Now, we have:

$$\ln \rho(x) = \ln(\phi(x)) - \ln(\bar{\Phi}(x)) = -\ln(\sqrt{2\pi}) - \frac{x^2}{2} - \ln(\bar{\Phi}(x))$$

Therefore

$$\frac{d}{dx} \ln \rho(x) = -x - \frac{1}{\bar{\Phi}(x)} \frac{d}{dx} \bar{\Phi}(x) = \frac{\phi(x)}{\bar{\Phi}(x)} - x = \rho(x) - x$$

But we also have $\frac{d}{dx}(-\ln \rho(x)) = \frac{d}{dx} \ln(1/\rho(x))$ so

$$\begin{aligned} \frac{d}{dx}(\rho(x) - x) &= -\frac{d^2}{dx^2} \ln \frac{1}{\rho(x)} \\ &= -\frac{d}{dx} \left(\rho(x) \left(\frac{d}{dx} \frac{1}{\rho(x)} \right) \right) \\ &= -\left(\rho(x) \left(\frac{d^2}{dx^2} \frac{1}{\rho(x)} \right) + \left(\left(\frac{d}{dx} \rho(x) \right) \left(\frac{d}{dx} \frac{1}{\rho(x)} \right) \right) \right) \end{aligned}$$

and

$$\frac{d}{dx} \frac{1}{\rho(x)} = -\frac{d}{dx} \rho(x) \frac{1}{\rho(x)^2} \implies \frac{d}{dx} \rho(x) = -\rho(x)^2 \left(\frac{d}{dx} \frac{1}{\rho(x)} \right)$$

Therefore

$$\frac{d}{dx}(\rho(x) - x) = -\left(\rho(x) \left(\frac{d^2}{dx^2} \frac{1}{\rho(x)} \right) - \rho(x)^2 \left(\frac{d}{dx} \frac{1}{\rho(x)} \right)^2 \right)$$

The term in the parenthesis in the variance of the measure $\nu_x(dz) := \rho(x)e^{-zx}\mu(dz)$, which is strictly positive. In fact for any $x \geq 0$, $\nu_x(dz)$ is a probability measure:

$$\int_0^\infty \nu_x(dz) = \rho(x) \int_0^\infty e^{-zx-z^2/2} dz = \rho(x) \int_x^\infty \frac{\phi(t)}{\phi(x)} dt = 1$$

Thus, $\rho(x) - x$ is decreasing because its first derivative is strictly negative. Now since

$$\frac{d^2}{dx^2} \ln \rho(x) = \frac{d}{dx}(\rho(x) - x) < 0$$

then the function $\ln \rho(x)$ is thus concave. To see that $\rho(x) - x$ vanishes as $x \rightarrow \infty$:

$$0 < \rho(x) - x < \frac{x}{x^2 - 1} \xrightarrow{x \rightarrow \infty} 0$$

It remains to show the last claim. For $\delta > 0$ and some $x^* \in (x, x + \delta)$, since $\rho(x) - x$ is decreasing:

$$\ln \frac{\rho(x + \delta)}{\rho(x)} = \delta \frac{\ln \rho(x + \delta) - \ln \rho(x)}{\delta} \stackrel{\text{MVT}}{=} \delta(\rho(x^*) - x^*) < \delta(\rho(x) - x)$$

and we conclude. □

Corollary 2.1. For $x \geq 0$ and $\delta > 0$:

$$e^{-\rho(x)\delta - \delta^2/2} \leq \frac{\bar{\Phi}(x + \delta)}{\bar{\Phi}(x)} \leq e^{-x\delta - \delta^2/2}$$

Proof. We have

$$\begin{aligned} \bar{\Phi}(x + \delta) &= \int_x^\infty \phi(z + \delta) dz \\ &= \frac{e^{-\delta^2/2}}{\sqrt{2\pi}} \int_x^\infty e^{-z^2/2 - z\delta} dz \\ &\stackrel{e^{-z\delta} \leq e^{-x\delta}, \forall z \geq x}{\leq} \frac{e^{-\delta^2/2 - x\delta}}{\sqrt{2\pi}} \int_x^\infty e^{-z^2/2} dz \\ &= e^{-\delta^2/2 - x\delta} \bar{\Phi}(x) \end{aligned}$$

By Theorem 2.2 we have $\ln \frac{\rho(x+\delta)}{\rho(x)} < (\rho(x) - x)\delta$. We also have

$$\begin{aligned} \frac{\bar{\Phi}(x + \delta)}{\bar{\Phi}(x)} &= \frac{\bar{\Phi}(x + \delta)}{\phi(x + \delta)} \frac{\phi(x)}{\bar{\Phi}(x)} \frac{\phi(x + \delta)}{\phi(x)} \\ &= \frac{\rho(x + \delta)}{\rho(x)} \frac{\phi(x + \delta)}{\phi(x)} \\ &= e^{-\ln \frac{\rho(x+\delta)}{\rho(x)}} e^{-\delta^2/2 - x\delta} \\ &\geq e^{-\rho(x)\delta - \delta^2/2} \end{aligned}$$

and we conclude. □

3 Laws of Iterated Logarithm

3.1 Motivation

Let X_1, X_2, \dots be a sequence of independently and identically distributed random variables with mean zero and variance one, and let $S_n = X_1 + \dots + X_n$ for $n \geq 1$ and $S_0 = 0$ be the associated partial sum process. The strong law of large numbers states that scaled by n^{-1} the partial sum process converges almost surely to the mean of the X 's, in this case zero:

$$n^{-1}S_n \xrightarrow{a.s.} 0 \text{ as } n \rightarrow \infty. \quad (4)$$

Whereas the central limit theorem states that scaled by $n^{-1/2}$ the partial sum process converges in distribution to a standard Gaussian:

$$n^{-1/2}S_n \xrightarrow{d} \mathcal{N}(0, 1) \text{ as } n \rightarrow \infty. \quad (5)$$

It is often of interest to quantify the fluctuations of S_n . The law of iterated logarithm provides such a result, and states that almost surely the fluctuations will be no larger than $\sqrt{2n \log \log(n)}$. Comparing to the scaling factors in (4) and (5) the law of iterated logarithm can be understood as operating *between* the central limit theorem and the law of large numbers.

3.2 A law of iterated logarithm for the Wiener process

We present a law of iterated logarithm for the Wiener process. In section 3.3 this result will be used as a building block for proving laws of iterated logarithms for more general random walks.

Theorem 3.1. *Let $(B_t)_{t \geq 0}$ be a Wiener process, then $\limsup_{n \rightarrow \infty} B_n / \sqrt{2n \log \log(n)} = 1$, \mathbb{P} -almost surely.*

We present the proof of Theorem 3.1 which can be found in section 8.5 of Durrett (2019). The proof relies on the reflection principle for Wiener processes, which is stated below.

Lemma 3.1. *Let $(B_t)_{t \geq 0}$ be a Wiener process, putting $T_a = \inf \{t > 0 \mid B_t = a\}$ with $a > 0$ it holds that $\mathbb{P}(T_a < t) = 2\mathbb{P}(B_t \geq a)$.*

Proof. If B hits a at some time $s < t$ then $B_t - B_{T_a}$ will be independent of $B_{T_a^-}$. The symmetry of the normal distribution along with the fact that $\mathbb{P}_a(B_u = a) = 0$ for $u > 0$ then imply that

$$\mathbb{P}(T_a < t, B_t > a) = \frac{1}{2} \mathbb{P}(T_a < t). \quad (6)$$

□

Above we have written \mathbb{P}_x for the probability conditioned on the fact that the Wiener process starts from x . Strictly speaking (6) needs to be proved formally, and a proof can be found on page 323 of Durrett (2019). Using the fact that $\{B_t > a\} \subset \{T_a < t\}$ and rearranging (6) gives the desired result. With this result in place we are now able to prove Theorem 3.1.

Proof. Using Lemma 3.2 we have that

$$\mathbb{P}\left(\sup_{s \in [0,1]} B_s > x\right) = \mathbb{P}(T_x \leq 1) = 2\bar{\Phi}(x) \quad (7)$$

Lemma 2.1 implies that $\bar{\Phi}(x) \sim \frac{1}{\sqrt{2\pi}} x^{-1} e^{-x^2/2}$ as $x \rightarrow \infty$, therefore combining this with the scaling property of Wiener processes we have that for any increasing function f

$$\mathbb{P}\left(B_n > (nf(n))^{1/2}\right) \sim \frac{1}{\sqrt{2\pi}} f(n)^{-1/2} \exp(-f(n)/2). \quad (8)$$

Using (8) we therefore have for any $\varepsilon > 0$ that

$$\sum_{n=1}^{\infty} \mathbb{P} \left(B_n > (nf(n))^{1/2} \right) \begin{cases} < \infty & \text{when } f(n) = (2 + \varepsilon) \log(n) \\ = \infty & \text{when } f(n) = (2 - \varepsilon) \log(n) \end{cases}. \quad (9)$$

Hence, the Borel-Cantelli lemma gives that \mathbb{P} -almost surely $\limsup_{n \rightarrow \infty} B_n / \sqrt{2n \log(n)} \leq 1$. To replace the $\log(n)$ term with $\log \log(n)$ we need to look at an exponentially growing sequence. Let $t_n = \alpha^n$ for some $\alpha > 1$. Then

$$\mathbb{P} \left(\sup_{s \in [t_n, t_{n+1}]} B_s > (t_n f(t_n))^{1/2} \right) \leq \mathbb{P} \left(\sup_{s \in [t_n, t_{n+1}]} B_s / \sqrt{t_{n+1}} > \left(\frac{f(t_n)}{\alpha} \right)^{1/2} \right) \quad (10)$$

$$\leq \frac{1}{\sqrt{2\pi}} \left(\frac{f(t_n)}{\alpha} \right)^{-1/2} \exp(-f(t_n)/2\alpha). \quad (11)$$

If $f(n) = 2\alpha^2 \log \log(n)$ then $f(t_n) = \log(n) + \log \log(\alpha)$ and therefor $\exp(-f(t_n)/2\alpha) \leq C_\alpha n^{-\alpha}$ for some $C_\alpha > 0$ depending only on α . Hence

$$\sum_{n=1}^{\infty} \mathbb{P} \left(\sup_{s \in [t_n, t_{n+1}]} B_s > (t_n f(t_n))^{1/2} \right) < \infty. \quad (12)$$

Since $t \mapsto (tf(t))^{1/2}$ is increasing and α is arbitrary we obtain that \mathbb{P} -almost surely

$$\limsup_{n \rightarrow \infty} B_n / \sqrt{2 \log \log(n)} \leq 1. \quad (13)$$

To prove a weak inequality in the other direction let t_n be defined as before but this time take α to be large in order to get independence. We put $\beta = t_{n+1}/(t_{n+1} - t_n) = \alpha/(\alpha - 1)$ and look at the exceedance probability of the increments:

$$\mathbb{P} \left(B_{t_{n+1}} - B_{t_n} > (t_{n+1} f(t_{n+1}))^{1/2} \right) = \mathbb{P} \left(B_1 > (\beta f(t_{n+1}))^{1/2} \right). \quad (14)$$

Again using Lemma 2.1 (14) will be no smaller than $\frac{1}{2\sqrt{2\pi}} (\beta f(t_{n+1}))^{-1/2} \exp(-\beta f(t_{n+1})/2)$.

Therefore, for n sufficiently large and putting $f(n) = (2/\beta^2) \log \log(n)$ we obtain that $\exp(-\beta f(t_{n+1})/2)$ is no smaller than $C_\alpha n^{-1/\beta}$. Therefore:

$$\sum_{n=0}^{\infty} \mathbb{P} \left(B_{t_{n+1}} - B_{t_n} > (t_{n+1} f(t_{n+1}))^{1/2} \right) = \infty. \quad (15)$$

Since the events in question are independent we obtain by Borel-Cantelli that infinity often

$$B_{t_{n+1}} - B_{t_n} > \sqrt{(2/\beta^2) t_{n+1} \log \log(t_{n+1})}. \quad (16)$$

Plugging in the definition of β , and using the facts that $t_n = t_{n+1}/\alpha$ along with the fact that $\log \log(n)$ is monotone increasing we obtain that

$$\limsup_{n \rightarrow \infty} B_{t_{n+1}} / \sqrt{2t_{n+1} \log \log(t_{n+1})} \geq \frac{\alpha - 1}{\alpha} - \frac{1}{\sqrt{\alpha}}. \quad (17)$$

Finally, letting $\alpha \rightarrow \infty$ proves the result. \square

3.3 Laws of iterated logarithm for discrete random walks

We now present a law of iterated logarithms for discrete random walks with finite second moment. The main theoretical device will be the embedding of random walks from independent random variables with bounded moments into Wiener processes. Such a result is given below.

Lemma 3.2. *Let X_1, X_2, \dots be independently distributed with mean zero, variance one, and common marginal distribution F . Putting $S_n = X_1 + \dots + X_n$ for $n \geq 1$ there is a sequence of stopping times T_0, T_1, \dots with $T_0 = 0$ such that $S_n \stackrel{d}{=} B_{T_n}$ and moreover the increments $T_n - T_{n-1}$ are independently distributed.*

The lemma is not hard to prove, and its proof can be found for example in Theorem 8.1.2 of Durrett (2019). However, we defer the proof to Chapter 10 where couplings of empirical processes and Wiener process will be discussed in detail. Combining Theorem 3.1 with Lemma 3.2 readily yields the following result.

Theorem 3.2. *If X_1, X_2, \dots are independently distributed with mean zero, variance one, and common distribution function F then \mathbb{P} -almost surely $\limsup_{n \rightarrow \infty} S_n / \sqrt{2n \log \log(n)} = 1$.*

Another approach to proving the theorem is given in Chapter 11 of Pollard (2002)

3.4 Distributional convergence

It can be of interest to understand the distribution of the maximal fluctuations in a standardised random walk. The following result can be found for example in Kabluchko (2007).

Theorem 3.3. *Let $(B_t)_{t \geq 0}$ be a Wiener process and for $n > 1$ put $M_n = \sup_{s \in [0, n]} s^{-1/2} B_s$. Then for every fixed $z \in \mathbb{R}$ it holds that $\mathbb{P}(M_n \leq a_n + zb_n) \rightarrow \exp(e^{-z})$ as $n \rightarrow \infty$, where in particular the normalizing constants are given by*

$$a_n = \sqrt{2 \log \log(n)} + \frac{1/2 \log \log \log(n) - \log(2\sqrt{\pi})}{\sqrt{2 \log \log(n)}} \quad (18)$$

$$b_n = \frac{1}{\sqrt{2 \log \log(n)}}. \quad (19)$$

Using the same approach as in Section 3.3 Darling & Erdős (1956) used the above theorem to prove the following result.

Theorem 3.4. *Let X_1, X_2, \dots be independently distributed with marginal $\mathcal{N}(0, 1)$ distribution, put $S_n = X_1 + \dots + X_n$, and put $M_n = \max_{s \in \{1, \dots, n\}} s^{-1/2} S_s$. Then for any fixed $z \in \mathbb{R}$ it holds that $\mathbb{P}(M_n \leq a_n + zb_n) \rightarrow \exp(e^{-z})$ as $n \rightarrow \infty$, where the normalizing constants are the same as those given in Theorem 3.3.*

3.5 Non-asymptotic laws of iterated logarithm

In applications it often useful to have a finite sample counterpart to Theorems 3.1 and 3.2. To this end Balsubramani (2014) proves the following result.

Theorem 3.5. *Let X_1, X_2, \dots be mutually independent Rademacher random variables and put $S_n = X_1 + \dots + X_n$. There is an absolute constant C , where it suffices to take $C = 137$, such that for any $\delta \in (0, 1)$ and any $n > C \log(4/\delta)$ it holds with probability at least $1 - \delta$ that*

$$|S_n| \leq \sqrt{3n \left(2 \log \log \left(\frac{5}{2} n \right) + \log \left(\frac{2}{\delta} \right) \right)} \vee 1. \quad (20)$$

The leading constant $\sqrt{6}$ in (20) does not seem optimal, as in fact Theorem 3.1 suggests the constant should be $\sqrt{2}$. Nevertheless, the result captures the first order asymptotics of the maximum fluctuations in the random walk, and makes explicit the trade off between δ and n . Balsubramani (2014) extend the result to martingales with uniformly bounded increments. Using the strong approximation results in Komlós et al. (1975) it should be possible to extend Theorem 3.5 to other discrete random walks not necessarily having bounded increments.

References

- Balsubramani, A. (2014), ‘Sharp finite-time iterated-logarithm martingale concentration’, *arXiv preprint arXiv:1405.2639*.
- Darling, D. A. & Erdős, P. (1956), ‘A limit theorem for the maximum of normalized sums of independent random variables’.
- Durrett, R. (2019), *Probability: theory and examples*, Vol. 49, Cambridge university press.
- Kabluchko, Z. (2007), ‘Extreme-value analysis of standardized gaussian increments’, *arXiv preprint arXiv:0706.1849*.
- Komlós, J., Major, P. & Tusnády, G. (1975), ‘An approximation of partial sums of independent rv’s, and the sample df. i’, *Zeitschrift für Wahrscheinlichkeitstheorie und verwandte Gebiete* **32**, 111–131.
- Pollard, D. (2002), *A user’s guide to measure theoretic probability*, number 8, Cambridge University Press.