Useful Inequalities in Probability and Statistics

Pietro Maria Sparago[a]

[a]Department of Statistics, London School of Economics and Political Science, London, UK

ARTICLE HISTORY
Compiled March 26, 2024

1. Tails of the normal distribution

Theorem 1.1. *Let $X \sim \mathcal{N}(0,1)$. Define $\overline{\Phi}(x) := P(X > x), x \geq 0$. Then*

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

Proof. We have, by continuity and Corollary 1.4:

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

We conclude for the upper bound with Lemma 1.2. □

Remark 1. Note that Proposition 1.1 yields a stronger version of the infimum Chernoff bound which is found with

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

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

1.1. Lemmas for the proof of Theorem 1.1 - mostly from [1]; Appendix D

Lemma 1.2. *Let $\phi(x) := (2\pi)^{-1/2}e^{-x^2/2}$. Then:*

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

Proof. We have $\overline{\Phi}(x) = \int_x^\infty \phi(t)dt$. 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. □

CONTACT P. M. Author. Email: p.sparago@lse.ac.uk

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

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

Theorem 1.3. *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

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

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 $\lim_{x\to-\infty} \frac{1}{\rho(x)} = \infty$. 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:

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

Therefore

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

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 $x\to\infty$:

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

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

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

and we conclude. $\square$

Corollary 1.4. *For $x \geq 0$ and $\delta > 0$:*

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

Proof. We have

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

By Theorem 1.3 we have $\ln \frac{\rho(x+\delta)}{\rho(x)} < (\rho(x) - x)\delta$. We also have

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

and we conclude. $\square$

1.2. Law of iterated logarithm - mostly from [1], Chapters 11.1 and 11.2

Lemma 1.5. *Let $S_n = X_1 + ... + X_n$ for $(X_n)_{n\leq N}$ IID $\sim \mathcal{N}(0,1)$. Then for $x \geq 0$*

$$P(\max_{n\leq N} S_n \geq x) \leq 2P(S_N \geq x)$$

Proof. We have $P(S_N - S_n \geq 0) = 1/2$ for any $n < N$ by the fact that $S_N - S_n \sim \mathcal{N}(0, N-n)$. Define $\tau = \inf\{n \leq N : S_n \geq x\} \wedge N$. The events $\{\tau = n\}, n \in \{1, ..., N\}$ are disjoint. So we have

$$\begin{aligned}
P(\max_{n\leq N} S_n \geq x) &= P(\cup_{n\leq N}\{\tau = n\}) \\
&= P(\cup_{n\leq N}(\{\tau = n\} \cap \{S_n \geq x\})) \\
&= \sum_{n\leq N} P(\{\tau = n\} \cap \{S_n \geq x\}) \\
&= 2\sum_{n<N} P(\{\tau = n\} \cap \{S_n \geq x\})P(S_N - S_n \geq 0) \\
&+ P(\{\tau = N\} \cap \{S_N \geq x\}) \\
&\overset{(*)}{=} 2\sum_{n<N} P(\{\tau = n\} \cap \{S_n \geq x\} \cap \{S_N - S_n \geq 0\}) \\
&+ P(\{\tau = N\} \cap \{S_N \geq x\}) \\
&\overset{(**)}{\leq} 2\sum_{n\leq N} P(\{\tau = n\} \cap \{S_N \geq x\}) \\
&= 2P(S_N \geq x)
\end{aligned}$$

where equality $(*)$ follows from the fact that $S_N - S_n$ is independent of S_n, while the inequality $(**)$ follows from the fact that $\{S_n \geq x\} \cap \{S_N - S_n \geq 0\} \subseteq \{S_N \geq x\}$ and $P(\{\tau = N\} \cap \{S_N \geq x\}) \leq 2P(\{\tau = N\} \cap \{S_N \geq x\})$. $\square$

Theorem 1.6. *Let $S_n = X_1 + ... + X_n$ where $(X_n)_{n\in\mathbb{N}}$ are IID $\sim \mathcal{N}(0,1)$. Then:*

(1) $\limsup_{n\to\infty} \frac{S_n}{\sqrt{2n\ln\ln n}} = 1$ *a.s.;*
(2) $\liminf_{n\to\infty} \frac{S_n}{\sqrt{2n\ln\ln n}} = -1$ *a.s.;*
(3) $\frac{S_n}{\sqrt{2n\ln\ln n}} \in B$ *i.o. a.s. for any open* $B \subseteq [-1,1]$.

Proof. (i). Let B be a block of consecutive positive integers (i.e. $B = \{n+1, ..., n+k\}$ for some n, k). Let $\gamma > 1$. Then by Lemma 1.5 and Theorem 1.2

$$\begin{aligned}
P(\cup_{\ell \in B}\{S_\ell \geq \gamma\sqrt{2\ell \ln\ln \ell}\}) &\leq P(\max_{\ell \in B} S_\ell \geq \gamma\sqrt{2n\ln\ln n}) \\
&\leq 2P(S_{n+k} \geq \gamma\sqrt{2n\ln\ln n}) \\
&= 2P\left(\frac{S_{n+k}}{\sqrt{n+k}} \geq \gamma\frac{\sqrt{2n\ln\ln n}}{\sqrt{n+k}}\right) \\
&= 2\overline{\Phi}\left(\gamma\frac{\sqrt{2n\ln\ln n}}{\sqrt{n+k}}\right) \\
&\leq \exp\left(-\frac{1}{2}\gamma^2\frac{2n\ln\ln n}{n+k}\right)
\end{aligned}$$

Now consider a sequence of consecutive blocks $B_k = \{n : n_k < n \leq n_{k+1}\}$ so that $n_k/\rho^k \to 1$ for some constant $\rho \in (1, \gamma)$. By the above then we then have the bounds

$$\begin{aligned}
P(\cup_{\ell \in B_k}\{S_\ell \geq \gamma\sqrt{2\ell\ln\ln \ell}\}) &\leq \exp\left(-\frac{1}{2}\rho^2\frac{2n_k\ln\ln n_k}{n_{k+1}}\right) \\
&\overset{k \text{ large enough}}{\approx} \exp\left(-\frac{1}{2}\rho^2\frac{2\rho^k\ln\ln\rho^k}{\rho^{k+1}}\right) \\
&= \exp\Big(-\rho(\ln k + \ln\ln\rho)\Big) \\
&= C(\rho)k^{-\rho}
\end{aligned}$$

which implies, by Borel-Cantelli I, that $P(S_n \geq \gamma\sqrt{2n\ln\ln n} \text{ i.o.}) = 0$. Therefore $P(\limsup_{n\to\infty}\frac{S_n}{\sqrt{2n\ln\ln n}} \leq \gamma) = 1$ for any $\gamma > 1$. We conclude that, considering a sequence $\gamma_\ell \downarrow 1$,

$$P\left(\limsup_{n\to\infty}\frac{S_n}{\sqrt{2n\ln\ln n}} \leq 1\right) = 1$$

We now need to prove that $\limsup_{n\to\infty}\frac{S_n}{\sqrt{2n\ln\ln n}} \geq 1$

$\square$

References

[1] Pollard, D. (2001). *A User's Guide to Measure Theoretic Probability*. Cambridge University Press.