EXACT BEHAVIOR OF GAUSSIAN SEMINORMS

Amir Dembo[1], Eddy Mayer–Wolf[2] and Ofer Zeitouni[3]

12 April 1993

ABSTRACT

The exact lower tail of Gaussian seminorms are evaluated, using a refinement of the techniques presented in [5].

KEYWORDS Gaussian norms. Large Deviations.

AMS 1991 Subject Classification Primary 60G15, 60F10. Secondary 60F05.

I. Introduction Let $\{x_i\}_{i=1}^{\infty}$ be a sequence of independent, standard Gaussian random variables. Consider the random variable $z = \sum_{i=1}^{\infty} x_i^2/a_i^2$, where $\{a_i\}_{i=1}^{\infty}$ is a sequence of given (deterministic) finite numbers satisfying $\sum_{i=1}^{\infty} \frac{1}{a_i^2} < \infty$. We are interested in computing the asymptotics $P(z < \epsilon)$ as $\epsilon \to 0$. This problem was considered in the 70's by the Soviet school [6, 7, 3]. Their approach was analytical in flavor and based on the saddle point method. Independently, unaware of this, a probabilistic study of $P(z < \epsilon)$ was initiated by [2] and improved upon in [5], the latter using large deviations techniques. This study provides explicit and easy bounds on $P(z < \epsilon)$ which, however, are not asymptotically tight. The purpose of this note is to push this analysis to yield a tight asymptotic evaluation of the above probability. We thus retrieve, by probabilistic methods, the analytic results of [6, 7, 3], in a shorter, more transparent way, which seems amenable to extensions in different directions.

To state precisely our goal, recall that the results in [5] imply that there exist functions

[1]Department of Mathematics and Department of Statistics, Stanford University, Stanford, CA 94305.
Partially supported by NSF DMS92-09712 grant and by a US-ISRAEL BSF grant.

[2]Department of Mathematics, Technion- Israel Institute of Technology, Haifa 32000, ISRAEL. Partially supported by the fund for promotion of research at the Technion.

[3]Department of Electrical Engineering, Technion- Israel Institute of Technology, Haifa 32000, ISRAEL.
Partially supported by a US-Israel BSF grant and by the fund for promotion of research at the Technion.

$g(\cdot)$ and $I(\cdot)$ such that z/ϵ satisfies the large deviations principle with rate $g(\epsilon)$ and rate function $I(\cdot)$. In particular, it is shown there that for some constant $c > 0$, which depends on the sequence $\{a_i\}$ and may be computed explicitly,

$$\lim_{\epsilon\to 0} g(\epsilon)\log P(z<\epsilon) = -c\,. \tag{1}$$

We show below that one may find an explicit function $f(\epsilon)$ such that

$$\lim_{\epsilon\to 0} f(\epsilon)P(z<\epsilon) = 1 \tag{2}$$

(see (10)). For related results and comparison theorems for small balls probabilities see [4].

Acknowledgment We thank Wenbo Li for bringing Sytaya's work [6] to our attention.

II. Computation of Asymptotic Probabilities As in the Introduction, let $\{x_i\}$ be i.i.d. standard Gaussian variables, defined on a suitable probability space $(\Omega,\mathcal{F},P)$. For $\theta \geq 0$ let $\Lambda(\theta) \triangleq \log E(e^{-\theta x_1^2}) = -\frac{1}{2}\log(1+2\theta)$ and define

$$\mu(\theta) \triangleq -\sum_{i=1}^{\infty}\frac{1}{a_i^2}\Lambda'(\frac{\theta}{a_i^2}) = \sum_{i=1}^{\infty}\frac{1}{a_i^2+2\theta} \tag{3}$$

$$\psi(\theta) \triangleq \sqrt{\sum_{i=1}^{\infty}\frac{\theta^2}{a_i^4}\Lambda''(\frac{\theta}{a_i^2})} = \sqrt{\sum_{i=1}^{\infty}\frac{2\theta^2}{(a_i^2+2\theta)^2}} \tag{4}$$

$$I(\theta) \triangleq \sum_{i=1}^{\infty}\left(\frac{\theta}{a_i^2}\Lambda'(\frac{\theta}{a_i^2}) - \Lambda(\frac{\theta}{a_i^2})\right) = \frac{1}{2}\sum_{i=1}^{\infty}\log(1+\frac{2\theta}{a_i^2}) - \theta\mu(\theta) \tag{5}$$

Under our assumptions, there exists a unique θ_ϵ such that $\mu(\theta_\epsilon) = \epsilon$, $\theta_\epsilon \to_{\epsilon\to 0} \infty$, and $\psi(\theta_\epsilon) \to_{\epsilon\to 0} \infty$. In the sequel, we use θ,μ,ψ,I for $\theta_\epsilon,\mu(\theta_\epsilon),\psi(\theta_\epsilon),I(\theta_\epsilon)$. Define on $(\Omega,\mathcal{F})$ a probability measure by

$$dP_\theta = \frac{e^{-\theta z}dP}{E_P(e^{-\theta z})} \tag{6}$$

Then

$$P(z<\epsilon) = e^{-I}\int_\Omega e^{\theta(z-\mu)}1_{\{0\leq z<\mu\}}dP_\theta \tag{7}$$

Defining now

$$U_\psi = \frac{\theta(\mu - z)}{\psi},$$

and denoting by η_ψ the law of U_ψ under P_θ, one obtains

$$e^I P(z<\epsilon) = \int_\Omega e^{-\psi U_\psi}1_{\{0<U_\psi\leq\theta\mu/\psi\}}dP_\theta = \int_0^{\mu\theta/\psi} e^{-\psi v}d\eta_\psi(v) = \int_0^\infty e^{-\psi v}d\eta_\psi(v) \tag{8}$$

where we used the fact that η_ψ is supported on $(-\infty, \theta\mu/\psi]$. Integrating by parts, one obtains

$$e^I P(z < \epsilon) = \int_0^\infty e^{-v} \eta_\psi([0, v/\psi]) dv \,. \tag{9}$$

It is straightforward to check that $E_{\eta_\psi}(U_\psi) = 0$ and $E_{\eta_\psi}(U_\psi^2) = 1$. Thus, if we knew that for $\epsilon \to 0$, η_ψ converges locally (in the scale $1/\psi$) to the standard Gaussian law, one would conclude from (9) the following

Theorem:

$$\lim_{\epsilon \to 0} e^I \psi P(z < \epsilon) = 1/\sqrt{2\pi} \,. \tag{10}$$

Remark: The behavior of I in ϵ may be read off its definition, and in many cases can be made explicit (see, e.g., displays (16) and (17) in [5]). A similar computation yields also the behavior of ψ in ϵ.

Proof: Define the (infinite) triangular array

$$U_\psi = \sum_{k=1}^{\infty} b_k \xi_k, \tag{11}$$

where under η_ψ, $\{\xi_k\}_{k=1}^\infty$ is a $(1 - N(0,1)^2)$ i.i.d. sequence and

$$b_k = \frac{\theta}{\psi(a_k^2 + 2\theta)} \,. \tag{12}$$

We now claim that the laws η_ψ converge to the standard Gaussian distribution as $\psi \to \infty$. Indeed, since in (11) $E_{\eta_\psi} \xi_k = 0 \;\; \forall k \in \mathbb{N}$ and $\sum_{k=1}^\infty b_k^2 E_{\eta_\psi} \xi_k^2 = 1$, one only needs to prove that, for some suitable $N(\epsilon)$,

$$\lim_{\psi \to \infty} \left(\sum_{k=1}^{N(\epsilon)} b_k^3 E_{\eta_\psi} |\xi_k|^3 + E_{\eta_\psi} \left| \sum_{k=N(\epsilon)+1}^{\infty} b_k \xi_k \right|^3 \right) = 0 \,, \tag{13}$$

in order to apply classical CLT results for finite triangular arrays such as for example [1, Theorem 7.1.2] . (The condition (13) follows from expressing (11) as a finite sum). Note (it follows directly from (12)) that

$$\sum_{k=1}^{\infty} b_k^2 = \frac{1}{2} \,, \qquad\qquad b_k \leq \frac{1}{2\psi} . \tag{14}$$

Therefore, denoting $m_i = E\,|\xi_1|^i$,

$$\sum_{k=1}^{N(\epsilon)} b_k^3 E_{\eta_\psi} |\xi_k|^3 \leq m_3 \max_{k\in\mathbb{N}} b_k \sum_{k=1}^{\infty} b_k^2 \leq \frac{m_3}{4\psi} \overset{\psi\to\infty}{\longrightarrow} 0\,,$$

whereas

$$E_{\eta_\psi} \left| \sum_{k=N(\epsilon)+1}^{\infty} b_k \xi_k \right|^3 \leq \left(E_{\eta_\psi} \left| \sum_{k=N(\epsilon)+1}^{\infty} b_k \xi_k \right|^4 \right)^{4/3} \leq \left(\sqrt{3m_4} \sum_{k=N(\epsilon)+1}^{\infty} b_k^2 \right)^{8/3} .$$

By taking any $N(\epsilon) \to_{\epsilon\to 0} \infty$, (13) follows.

It is actually possible to obtain a stronger (uniformly local) convergence. Let U_ψ's characteristic function be denoted by $\phi_\psi(t) = E_{\eta_\psi} e^{itU_\psi}$.

Lemma: For each $\alpha > 0$,

$$|\phi_\psi(t)| \leq \frac{\alpha^{\alpha/2}}{|t|^\alpha} \wedge 1 \qquad \forall \psi > \alpha^{1/2},\ \forall t \in \mathbb{R}. \tag{15}$$

Proof of the Lemma: Using the representation (11), a direct calculation yields $\phi_\psi(t) = e^{i\frac{\theta}{\psi}\mu t} \prod_{k=1}^\infty (1 + 2ib_k t)^{-\frac{1}{2}}$ so that

$$|\phi_\psi(t)|^4 = \prod_{k=1}^{\infty} \frac{1}{1 + 4b_k^2 t^2} = \exp - \left\{ \sum_{k=1}^{\infty} \log(1 + 4b_k^2 t^2) \right\}. \tag{16}$$

We now use the inequality

$$\log(1 + \gamma t^2) \geq \gamma \log(1 + t^2) \quad \forall \gamma \in (0, 1),\ \forall t \in \mathbb{R}\,, \tag{17}$$

in (16). Given $\alpha > 0$, and if $\psi > \alpha^{1/2}$, then by (14) $\gamma_k = 4\alpha b_k^2 < 1$ for all $k \in \mathbb{N}$ and

$$\begin{aligned} |\phi_\psi(t)|^4 &\leq \exp - \left\{ \sum_{k=1}^{\infty} 4\alpha b_k^2 \log(1 + t^2/\alpha) \right\} = \exp - \{2\alpha \log(1 + t^2/\alpha)\} \\ &= \left(1 + \frac{t^2}{\alpha}\right)^{-2\alpha} \end{aligned}$$

so that indeed $|\phi_\psi(t)| \leq \alpha^{\alpha/2} |t|^{-\alpha}$ for all $\psi > \alpha^{1/2}$ and all $t \in \mathbb{R}$. $\square$

It then follows that

$$\lim_{\psi\to\infty} \int_{-\infty}^{\infty} \left| \phi_\psi(t) - e^{-\frac{t^2}{2}} \right| dt = 0. \tag{18}$$

Indeed, the pointwise convergence of the integrand in (18) to 0 is guaranteed by the CLT for η_ψ, and the L^1 limit results by dominated convergence which is made possible by the Lemma (applied with any $\alpha > 1$).

Denote next by Γ the standard Gaussian measure and by $g(x)$ its density. From (9)

$$\left|\psi e^I P(z<\epsilon) - \frac{1}{\sqrt{2\pi}}\right| \le \left|\psi \int_0^\infty e^{-v}\Gamma([0,\frac{v}{\psi}])dv - \frac{1}{\sqrt{2\pi}}\right| + \psi \int_0^\infty e^{-v}|(\Gamma-\eta_\psi)([0,\frac{v}{\psi}])|dv \triangleq J_1 + J_2.$$

Now, for some $0 < \bar{v}_\psi < v/\psi$, $J_1 = |\int_0^\infty (g(\bar{v}_\psi) - g(0))ve^{-v}dv| \stackrel{\psi\to\infty}{\longrightarrow} 0$ by dominated convergence, while, denoting by f_ψ the density of the law η_ψ, whose existence follows from (15),

$$J_2 \le \psi \int_0^\infty e^{-v}\left(\int_0^{v/\psi} |g(x) - f_\psi(x)|dx\right) dv \le \|g - f_\psi\|_\infty \int_0^\infty ve^{-v}dv$$

which by (18) coupled with the inverse Fourier transform, converges to zero as $\psi \to \infty$. $\square$

Remarks:

1) We mention the particular case $a_i = i^{\beta/2}$, $\beta > 1$. A computation similar to the one in [5](displays (19), (20)) reveals that

$$I = \frac{(\beta-1)\nu_\beta^{\frac{\beta}{\beta-1}}}{2\epsilon^{\frac{1}{\beta-1}}}(1+o(1)), \quad \psi^2 = \frac{\nu_\beta^{\frac{1}{\beta-1}}\kappa_\beta}{2\epsilon^{\frac{1}{\beta-1}}}(1+o(1)),$$

where $\nu_\beta = \int_0^\infty dy/(1+y^\beta) = \frac{\pi/\beta}{\sin(\pi/\beta)}$ and $\kappa_\beta = \int_0^\infty dy/(1+y^\beta)^2 = \frac{(\beta-1)/\beta^2}{\sin((\beta-1)\pi/\beta)}$.

2) Our interest in the lower tails of the random variable z is because it represents, by the Karhunen-Loève expansion, the L^2 norm of Gaussian processes. This is the reason for our considering x_i which are Gaussian random variables. The method, and perhaps the result as well, seem to depend quite heavily on the Gaussian assumption. For example, taking $a_i^2 = 2^i$ and x_i Bernoulli random variables, z turns out to be uniform on $[0,1]$, and the functions μ, ψ and I may be computed explicitly; it turns out that $\psi \to_{\epsilon\to 0} 1$ and the limit in (10) is e^{-1} and not $(2\pi)^{-1/2}$, suggesting that the central limit theorem fails in this case.

Nevertheless, under the following conditions on $\Lambda(\theta)$ and on the sequence $\{a_k^2\}$, the derivation may still be carried out, defining $\mu(\theta)$, $\psi(\theta)$ and $I(\theta)$ as in the first equality in each of (3)–(5):

(A1) $\lim_{\theta\to\infty}\Lambda'(\theta)=0$

(A2) $\lim_{\theta\to\infty}\psi(\theta)=\infty$

(A3) (a) $\Lambda''(y)\le c/y^2\ \forall y>0$.

(b) $|\Lambda^{(4)}(y)|\le c\Lambda''(y)^2\ \forall y>0$.

(A4) for m large enough, and analytically extending Λ to $\mathrm{Re}\theta>0$,

$$\Lambda(x)-\mathrm{Re}\Lambda(x+it\frac{x}{m})\ge(\frac{x}{m})^2\Lambda''(x)G(t)\ \forall x>0,t\in\mathbb{R}, \tag{19}$$

where $G(t)$ satisfies $\int_{-\infty}^{\infty}e^{-G(t)}dt<\infty$.

Note that **(A1), (A3)** and **(A4)** only involve x_1's law while **(A2)** depends also on the sequence $\{a_k^2\}$. A sufficient condition for **(A2)** to hold for any nondegenerate x_1 is that a_i^2 grows subexponentially in the sense that for some $0<a<b<\infty$ the cardinality of $\{i: a\theta\le a_i^2\le b\theta\}$ diverges with θ, since

$$\begin{aligned}\psi^2(\theta) &\ge \sum_{\{i:\frac{\theta}{a_i^2}\in[1/b,1/a]\}}\frac{\theta^2}{a_i^4}\Lambda''(\frac{\theta}{a_i^2})\\ &\ge m|\{i:\frac{1}{b}\le\frac{\theta}{a_i^2}\le\frac{1}{a}\}|=m|\{i:\ a\theta\le a_i^2\le b\theta\}|,\end{aligned}$$

where m denotes the positive minimum of the function $x^2\Lambda''(x)$ on the interval $[1/b,1/a]$.

The conditions are satisfied, for example, when x_1 has a density of the form $C_\alpha|x|^\alpha e^{-x^2/2}$, for some $\alpha>-1$, since the corresponding $\Lambda_\alpha(\theta)$ satisfy $\Lambda_\alpha=(1+\alpha)\Lambda_0$. On the other hand, if x_1^2 is a lattice random variable, the left hand side of (19) is periodic in t and thus **(A4)** cannot possibly hold. This is the situation, in particular, in the uniform example discussed above.

To see that these assumptions enable the above proof to be emulated, first note that **(A1)** implies, by the monotonicity of $\Lambda'(\cdot)$, that $\mu(\theta)\to_{\theta\to\infty}0$ and hence the existence of θ_ϵ is ensured. Next, U_ψ is represented as $U_\psi=\sum_{k=1}^{\infty}\overline{\xi}_k$ with $E_{\eta_\psi}\overline{\xi}_k=0$ and $\sum_{k=1}^{\infty}E_{\eta_\psi}\overline{\xi}_k^2=1$, the third moment condition in (13) is replaced by Hölder's inequality to a fourth moment condition, which in turn is handled using **(A2)** and **(A3)**. Finally, the uniform integrability follows from the Lemma as before, with **(A4)** replacing (17) and providing the required handle on $|\phi_\psi(t)|$.

References

[1] K. L. Chung. *A Course in Probability Theory.* Academic Press, New York, second edition, 1974.

[2] J. Hoffman-Jorgensen, L.A. Shepp, and R.M Dudley. On the lower tail of Gaussian seminorms. *Annals of Probability*, 7:319–342, 1979.

[3] I. A. Ibragimov. Hitting probability of a Gaussian vector with values in a Hilbert space in a sphere of small radius. *J. Sov. Math*, 20:2164–2174, 1982.

[4] W. V. Li. Comparison results for the lower tail of Gaussian seminorms. *J. Theor. Prob.*, 5:1–31, 1992.

[5] E. Mayer-Wolf and O. Zeitouni. The probability of small Gaussian ellipsoids and assoiciated conditional moments. *Annals of Probability*, 21, 1992.

[6] G. N. Sytaya. On some asymptotic representations of the gaussian measure in a Hilbert space. In *Theory of Stochastic processes, Publication 2*, pages 93–104. Ukrainan Academy of Science, 1974.

[7] V. M. Zolotarev. Asymptotic behavior of the Gaussian measure in l_2. *J. Sov. Math*, 24:2330–2334, 1986.