Tools of modern probability theory

Imre Péter Tóth, Math Inst. of TU Budapest

September 16, 2020

Abstract

These are lecture notes for some of the course "Tools of Modern Probability" for 3rd year BSc and 1st year MSc students of Mathematics at TU Budapest, in the autumn semester of 2020/21. Draft under construction, with many typos, errors and inconsistencies.

1 Gaussian integrals

Question:

$$I := \int_{-\infty}^{\infty} e^{-\frac{x^2}{2}} \mathrm{d}\, x = ?$$

Or, more generally:

$$\int_{-\infty}^{\infty} e^{-a(x-m)^2} \mathrm{d}\, x = ?$$

whenever $m \in \mathbb{R}$ and $0 < a \in \mathbb{R}$.

Difficulty: the indefinite integral

$$\Phi(x) := \int e^{-\frac{x^2}{2}} \mathrm{d}\, x$$

can not be expressed with elementary functions.

Solution: we calculate the double integral

$$V := \iint_{\mathbb{R}^2} e^{-\frac{x^2+y^2}{2}} \mathrm{d}\, x \mathrm{d}\, y$$

in two different ways:

1.) Notice that $e^{-\frac{x^2+y^2}{2}} = e^{-\frac{x^2}{2}} \cdot e^{-\frac{y^2}{2}}$ is a product of an x-dependent and a y-dependent function, so

$$V = \iint_{\mathbb{R}^2} e^{-\frac{x^2+y^2}{2}} \mathrm{d}\, x \mathrm{d}\, y = \int_{-\infty}^{\infty} \int_{-\infty}^{\infty} e^{-\frac{x^2}{2}} e^{-\frac{y^2}{2}} \mathrm{d}\, y \mathrm{d}\, x = \int_{-\infty}^{\infty} e^{-\frac{x^2}{2}} \mathrm{d}\, x \cdot \int_{-\infty}^{\infty} e^{-\frac{y^2}{2}} \mathrm{d}\, y = I^2.$$

2.) Notice that $x^2 + y^2 = r^2$ when we use polar coordinates – i.e. the function $e^{-\frac{x^2+y^2}{2}}$ is rotation symmetric. Using polar coordinates, $\mathrm{d}\, x \mathrm{d}\, y = r \mathrm{d}\, \varphi \mathrm{d}\, r$, so

$$V = \iint_{\mathbb{R}^2} e^{-\frac{x^2+y^2}{2}} \mathrm{d}\, x \mathrm{d}\, y = \int_{0}^{\infty} \int_{0}^{2\pi} e^{-\frac{r^2}{2}} r \mathrm{d}\, \varphi \mathrm{d}\, r = \int_{0}^{\infty} 2\pi r e^{-\frac{r^2}{2}} \mathrm{d}\, r = 2\pi \left[-e^{-\frac{r^2}{2}} \right]_0^{\infty} = 2\pi.$$

Comparing the two, we get that $I^2 = 2\pi$, so

$$I = \int_{-\infty}^{\infty} e^{-\frac{x^2}{2}} \mathrm{d}\,x = \sqrt{2\pi}.$$

Homework 1.1. *Use the integral substitution* $\frac{y^2}{2} := a(x-m)^2$ *to show that*

$$\int_{-\infty}^{\infty} e^{-a(x-m)^2}\, d\,x = \sqrt{\frac{\pi}{a}} \tag{1}$$

whenever $m \in \mathbb{R}$ *and* $0 < a \in \mathbb{R}$.

2 Polar coordinates in higher dimensions

In $\mathbb{R}^2$ the polar coordinate transformation

$$\begin{aligned} x &= r\cos\varphi \\ y &= r\sin\varphi \end{aligned}$$

has Jacobi matrix

$$J = \begin{pmatrix} \frac{\partial x}{\partial r} & \frac{\partial x}{\partial \varphi} \\ \frac{\partial y}{\partial r} & \frac{\partial y}{\partial \varphi} \end{pmatrix} = \begin{pmatrix} \cos\varphi & -r\sin\varphi \\ \sin\varphi & r\cos\varphi \end{pmatrix},$$

which has determinant

$$\det(J) = r\cos^2\varphi + r\sin^2\varphi = r.$$

So, in the integral transformation we have

$$\mathrm{d}\,x\mathrm{d}\,y = r\mathrm{d}\,\varphi\mathrm{d}\,r.$$

As a result, integrating rotation symmetric functions in $\mathbb{R}^2$ always boils down to a 1-dimensional integral in the following way: If $f : \mathbb{R}^2 \to \mathbb{R}$ and $f(x,y)$ depends on $r = \sqrt{x^2+y^2}$ only, e.g. $f(x,y) = \tilde{f}(r)$, then

$$\iint_{R^2} f(x,y)\mathrm{d}\,x\mathrm{d}\,y = \int_0^{\infty}\int_0^{2\pi} \tilde{f}(r) r\mathrm{d}\,\varphi\mathrm{d}\,r = \int_0^{\infty} 2\pi r\tilde{f}(r)\mathrm{d}\,r. \tag{2}$$

The factor $2\pi r$ is exactly the circumference of a circle with radius r. **Of course**: $2\pi r\mathrm{d}\,r$ is exactly the area of the annulus of width $\mathrm{d}\,r$ around the circle of radius r, if $\mathrm{d}\,r$ is infinitesimally small.

Similarly, for $d = 1, 2, 3, \ldots$ let $S^{d-1} \subset \mathbb{R}^d$ be the unit sphere in $\mathbb{R}^d$:

$$S^{d-1} := \left\{(x_1, \ldots, x_d) \in \mathbb{R}^d \middle| x_1^2 + x_2^2 + \cdots + x_d^2 = 1\right\}. \tag{3}$$

So, for example

- $S^0 = \{-1, 1\} \subset \mathbb{R}$ is a discrete set of only two points,
- S^1 is the unit circle in $\mathbb{R}^2$,
- S^2 is the usual unit sphere in 3D.

Let c_d denote the $d-1$-dimensional "surface volume" of S^{d-1}, so

- If $d=1$, then $d-1=0$, so c_1 is a "zero-dimensional size", meaning number of points.
- If $d=2$, then $d-1=1$, so c_2 is arclength.
- If $d=3$, then $d-1=2$, so c_3 is area.

In particular:

d	1	2	3
c_d	2	2π	4π

Important observation: If, instead of unit spheres, we look at the sphere of radius r in $\mathbb{R}^d$, the surface scales with r^{d-1} so that it is $c_d r^{d-1}$.

This proves (not very rigorously) the following generalization of (2):

Theorem 2.1. *If $f:\mathbb{R}^d\to\mathbb{R}$ and $f(x_1,\dots,x_d)=\tilde{f}(r)$ with $r=\sqrt{x_1^2+\cdots+x_d^2}$ (meaning that f is spherically symmetric), then*

$$\int_{\mathbb{R}^d} f(x_1,\dots,x_d)\,dx_1\dots dx_d = c_d\int_0^\infty r^{d-1}\tilde{f}(r)\,dr.$$

The following homework demonstrates a very clever way to calculate the numbers c_d using Gaussian integrals.

Homework 2.2. *Let $f(x_1,\dots,x_d)=e^{-\frac{x_1^2+\cdots+x_d^2}{2}}$, and let $V=\int_{\mathbb{R}^d} f(\underline{x})\,d\underline{x}$.*

- *Calculate V using that f is a product:*
$$f(x_1,\dots,x_d)=e^{-\frac{x_1^2}{2}}\cdot e^{-\frac{x_2^2}{2}}\cdot\dots\cdot e^{-\frac{x_d^2}{2}}.$$
- *Write V as a one-dimensional integral using Theorem 2.1.*
- *Compare the two results to get that*
$$c_d=\frac{\sqrt{2\pi}^d}{\int_0^\infty r^{d-1}e^{-\frac{r^2}{2}}\,dr}.$$

Remark 2.3. *The last integral is in some sense easy: the integral substitution $t=\frac{r^2}{2}$ (or $r=\sqrt{2t}$) gives $dr=\frac{1}{\sqrt{2t}}dt$, so*

$$\int_0^\infty r^{d-1}e^{-\frac{r^2}{2}}\,dr=\int_0^\infty \sqrt{2t}^{d-1}e^{-t}\frac{1}{\sqrt{2t}}dt=2^{\frac{d}{2}-1}\int_0^\infty t^{\frac{d}{2}-1}e^{-t}dt=2^{\frac{d}{2}-1}\Gamma\left(\frac{d}{2}\right),$$

where

$$\Gamma(s):=\int_0^\infty t^{s-1}e^{-t}\,dt$$

is the famous Euler gamma function, which we will discuss. Using this, we get

$$c_d=2\frac{\pi^{\frac{d}{2}}}{\Gamma\left(\frac{d}{2}\right)}.$$

Homework 2.4. *Let $B_d\subset\mathbb{R}^d$ be the unit ball is R^d meaning*

$$B_d:=\left\{(x_1,\dots,x_d)\in\mathbb{R}^d \middle| x_1^2+x_2^2+\cdots+x_d^2\le 1\right\}.$$

(Compare the definition (3) of the sphere.) Let b_d be the d-dimensional volume of B_d. Calculate b_d.

(Hint: let $f(r)$ be the surface of the sphere with radius r, and let $g(r)$ be the volume of the ball with radius r.) Convince me (and yourself) that $g'(r)=f(r)$.

3 Almost Gaussian integrals, Laplace's method

In this section we study integrals of the form $\int e^{nf(x)}$ where n is large.

Suppose that a function $f : \mathbb{R} \to \mathbb{R}$ has a unique absolute maximum at some x_0. Suppose for simplicity that $x_0 = 0$. So, everywhere else, f is less than at 0. Now let $n \in \mathbb{R}$ be big. Then the function $n \cdot f$ also has a unique maximum at 0, which is somehow more pronounced: differences in function values are magnified by the factor n. If we put all this into the exponent – meaning that we consider e^{nf}, then the unique maximum is still at 0, but choosing bigger and bigger n-s now magnifies the *ratio* of function values. Function values far away from 0 become "negligible", and the integral $\int_{\mathbb{R}} e^{nf(x)} \mathrm{d}\,x$ is dominated by the contribution of a small neighbourhood of 0.

Now if f is twice differentiable, then $f(x) \approx a_0 + a_1 x + a_2 x^2$ near x, where $a_0 = f(0)$ is the maximum value, $a_1 = f'(0) = 0$ because 0 is a critical point, and $a_2 = \frac{f''(0)}{2}$ is negative[1]. So let me use the notation $A := f(0)$ and $B := -f''(0)$ to write $e^{nf(x)} \approx e^{nA - \frac{nB}{2}x^2} = e^{nA} \cdot e^{-\frac{nB}{2}x^2}$ near 0, which means that

$$\int_{\mathbb{R}} e^{nf(x)} \mathrm{d}\,x \approx e^{nA} \int_{-\infty}^{\infty} e^{-\frac{nB}{2}x^2} \mathrm{d}\,x.$$

The right hand side is a Gaussian integral, and can be calculated explicitly using (1).

This argument is made precise in the following theorem:

Theorem 3.1. *Let $f : [a, b] \to \mathbb{R}$ be twice differentiable, and let $x_0 \in (a, b)$ be its unique maximum place. Suppose that $f''(x_0) < 0$. Then*

$$\int_a^b e^{nf(x)} d\,x \sim e^{nf(x_0)} \sqrt{\frac{2\pi}{n(-f"(x_0))}} \quad \textit{as } n \to \infty.$$

Remark 3.2. *If $a = -\infty$ and/or $b = \infty$, we need to make two additional assumptions:*

1. *Assume that not only is x_0 the unique global maximum place, but function values cannot even approach $f(x_0)$ as $x \to -\infty$ or $x \to \infty$. This can be formulated as $\limsup_{x\to-\infty} f(x) < f(x_0)$ and $\limsup_{x\to\infty} f(x) < f(x_0)$.*
2. *Assume that $\int_a^b e^{nf(x)} d\,x < \infty$ for some n.*

Proof. Without loss of generality, we can assume that $a < x_0 = 0 < b$. Let's use the notation $A := f(0)$ and $B := -f''(0)$. Notice that by (1), the statement of the theorem is equivalent to

$$\int_a^b e^{nf(x)} \mathrm{d}\,x \sim \int_{-\infty}^{\infty} e^{n\left(A - \frac{B}{2}x^2\right)} \mathrm{d}\,x = e^{nA} \sqrt{\frac{2\pi}{nB}} \quad \text{as } n \to \infty.$$

Before turning to the essence, we get rid of the minor difficulties:

Lemma 3.3. *For any $\delta > 0$*

$$\int_{-\delta}^{\delta} e^{n\left(A - \frac{B}{2}x^2\right)} d\,x \sim \int_{-\infty}^{\infty} e^{n\left(A - \frac{B}{2}x^2\right)} d\,x \quad \textit{as } n \to \infty,$$

[1] a_2 can be zero in the degenerate case, but let's assume it is not.

Proof. Using the integral substitution $x = \frac{y}{\sqrt{n}}$ we get

$$\int_{-\delta}^{\delta} e^{n\left(A-\frac{B}{2}x^2\right)} \mathrm{d}\, x = e^{nA} \int_{-\delta}^{\delta} e^{-\frac{nB}{2}x^2} \mathrm{d}\, x = e^{nA} \frac{1}{\sqrt{n}} \int_{-\sqrt{n}\delta}^{\sqrt{n}\delta} e^{-\frac{B}{2}y^2} \mathrm{d}\, y.$$

Since the improper integral $\int_{-\infty}^{\infty} e^{-\frac{B}{2}y^2} \mathrm{d}\, y$ is convergent,

$$\int_{-\sqrt{n}\delta}^{\sqrt{n}\delta} e^{-\frac{B}{2}y^2} \mathrm{d}\, y \to \int_{-\infty}^{\infty} e^{-\frac{B}{2}y^2} \mathrm{d}\, y \quad \text{as } n \to \infty,$$

so

$$\int_{-\delta}^{\delta} e^{n\left(A-\frac{B}{2}x^2\right)} \mathrm{d}\, x = \frac{e^{nA}}{\sqrt{n}} \int_{-\sqrt{n}\delta}^{\sqrt{n}\delta} e^{-\frac{B}{2}y^2} \mathrm{d}\, y \sim \frac{e^{nA}}{\sqrt{n}} \int_{-\infty}^{\infty} e^{-\frac{B}{2}y^2} \mathrm{d}\, y = \int_{-\infty}^{\infty} e^{n\left(A-\frac{B}{2}x^2\right)} \mathrm{d}\, x.$$

□

Lemma 3.4. *For any $\delta > 0$ such that $a < -\delta < \delta < b$*

$$\int_{-\delta}^{\delta} e^{nf(x)} dx \sim \int_{a}^{b} e^{nf(x)} dx \quad \textit{as } n \to \infty.$$

Proof. For $x \notin [-\delta, \delta]$ the function values are strictly bounded away from $A = f(0)$, so there is an $\eta > 0$ such that $f(x) \leq A - \eta$, and thus $e^{nf(x)} \leq e^{nA} e^{-n\eta}$ whenever $x \notin [-\delta, \delta]$. So

$$\int\limits_{[a-b]\setminus[-\delta,\delta]} e^{nf(x)} \mathrm{d}\, x \leq e^{nA}(b-a)e^{-n\eta}.$$

We will see in the next argument (fighting the main difficulty) that $\int_{-\delta}^{\delta} e^{nf(x)} \mathrm{d}\, x$ is in the order of magnitude of $\frac{e^{nA}}{\sqrt{n}}$, so the integral outside $[-\delta, \delta]$ is indeed negligible. □

Now we fight the main difficulty, which is the approximation of $e^{nf(x)}$ with Gaussians near $x_0 = 0$. Since f is twice differentiable at 0, Taylor's theorem says that

$$f(x) = A - \frac{B}{2}x^2 + o(x^2) \quad \text{as } x \to 0.$$

In particular, for any $\varepsilon > 0$ there is some $\delta > 0$ such that for all $|x| \leq \delta$

$$A - \frac{B}{2}x^2 - \frac{\varepsilon}{2}x^2 \leq f(x) \leq A - \frac{B}{2}x^2 + \frac{\varepsilon}{2}x^2,$$

which means

$$e^{nA} e^{-n\frac{B+\varepsilon}{2}x^2} \leq e^{nf(x)} \leq e^{nA} e^{-n\frac{B-\varepsilon}{2}x^2},$$

so

$$e^{nA} \int_{-\delta}^{\delta} e^{-n\frac{B+\varepsilon}{2}x^2} \mathrm{d}\, x \leq \int_{-\delta}^{\delta} e^{nf(x)} \mathrm{d}\, x \leq e^{nA} \int_{-\delta}^{\delta} e^{-n\frac{B-\varepsilon}{2}x^2} \mathrm{d}\, x.$$

We use Lemma 3.3 and (1) to calculate the asymptotics of both estimates:

$$e^{nA} \int_{-\delta}^{\delta} e^{-n\frac{B+\varepsilon}{2}x^2} \mathrm{d}\, x \quad \sim \quad e^{nA} \sqrt{\frac{2\pi}{n(B+\varepsilon)}} = e^{nA} \sqrt{\frac{2\pi}{nB}} \sqrt{\frac{B}{B+\varepsilon}}$$

$$e^{nA} \int_{-\delta}^{\delta} e^{-n\frac{B-\varepsilon}{2}x^2} \mathrm{d}\, x \quad \sim \quad e^{nA} \sqrt{\frac{2\pi}{n(B-\varepsilon)}} = e^{nA} \sqrt{\frac{2\pi}{nB}} \sqrt{\frac{B}{B-\varepsilon}}$$

This means that

$$\liminf_{n\to\infty} \frac{\int_{-\delta}^{\delta} e^{nf(x)} \mathrm{d}\, x}{e^{nA}\sqrt{\frac{2\pi}{nB}}} \geq \sqrt{\frac{B}{B+\varepsilon}},$$

$$\limsup_{n\to\infty} \frac{\int_{-\delta}^{\delta} e^{nf(x)} \mathrm{d}\, x}{e^{nA}\sqrt{\frac{2\pi}{nB}}} \leq \sqrt{\frac{B}{B-\varepsilon}}.$$

(This is the point when we are really done with the proof of Lemma 3.4. Now we apply it.) Using Lemma 3.4 we get rid of the δ-s in the boundary of the integral:

$$\liminf_{n\to\infty} \frac{\int_{a}^{b} e^{nf(x)} \mathrm{d}\, x}{e^{nA}\sqrt{\frac{2\pi}{nB}}} \geq \sqrt{\frac{B}{B+\varepsilon}},$$

$$\limsup_{n\to\infty} \frac{\int_{a}^{b} e^{nf(x)} \mathrm{d}\, x}{e^{nA}\sqrt{\frac{2\pi}{nB}}} \leq \sqrt{\frac{B}{B-\varepsilon}}.$$

Since this holds for every $\varepsilon > 0$,

$$\lim_{n\to\infty} \frac{\int_{a}^{b} e^{nf(x)} \mathrm{d}\, x}{e^{nA}\sqrt{\frac{2\pi}{nB}}} = 1,$$

so the theorem is proven. □

4 Euler gamma function

Definition 4.1. *For $s > 0$ let*

$$\Gamma(s) = \int_0^\infty x^{s-1} e^{-x} dx$$

be the Euler gamma function.

Remark 4.2. *The definiton can be exetended to all complex s except for the non-positive integers, but we don't do that now.*

Homework 4.3. *Check that $\Gamma(s+1) = s\Gamma(s)$ for all $s > 0$.*

Homework 4.4. *Check by induction that $\Gamma(n+1) = n!$ for all $n \in \mathbb{N}$.*

Homework 4.5. *Calculate $\Gamma\left(\frac{1}{2}\right)$.*

The gamma function is a beautiful generalization of the factorial for non-integer numbers. In fact, it is strictly convex on $(0, \infty)$. However, it is not monotonous: it has a minimum at $x \approx 1.46$.

Values of the gamma function at half-integers are of special interest, see Remark 2.3.

s	$\frac{1}{2}$	1	$\frac{3}{2}$	2	$\frac{5}{2}$	3	$\frac{7}{2}$	4
$\Gamma(s)$	$\sqrt{\pi}$	1	$\frac{1}{2}\sqrt{\pi}$	1	$\frac{3}{4}\sqrt{\pi}$	2	$\frac{15}{8}\sqrt{\pi}$	6

5 Stirling's approximation

Stirling's approximation is a great tool to do calculations with factorials. It has many versions, depending on the accuracy. Here we present a version which is easy to prove and accurate enough for our purposes.

Theorem 5.1 (Striling's approximation)**.**

$$\Gamma(n+1) \sim n^n e^{-n} \sqrt{n}\sqrt{2\pi} \quad \textit{as } n \to \infty.$$

In particular,

$$n! \sim n^n e^{-n} \sqrt{n}\sqrt{2\pi} \quad \textit{as } n \to \infty.$$

The first version, despite of the notation, is for $n \in \mathbb{R}$*, not only integers.*

Remark 5.2. *This formula is often written as* $n! \sim \sqrt{2\pi n}\left(\frac{n}{e}\right)^n$ *for compactness. However, I prefer the form as written in the theorem, because the terms there are in the order of significance.*

Proof. By definition, $\Gamma(n+1) = \int_0^\infty x^n e^{-x} \mathrm{d}\, x$. Using the integral substitution $x = ny$ we get

$$\Gamma(n+1) = e^{n \ln n} n \int_0^\infty e^{n(\ln y - y)} \mathrm{d}\, y.$$

We apply Theorem 3.1 to the function $f(y) = \ln y - y$. It has its unique maximum at $y_0 = 1$ with $f(1) = -1$ and $f''(1) = -1$. So Theorem 3.1 gives $\int_0^\infty e^{n(\ln y - y)} \mathrm{d}\, y \sim e^{-n}\sqrt{\frac{2\pi}{n}}$, which completes the proof. □