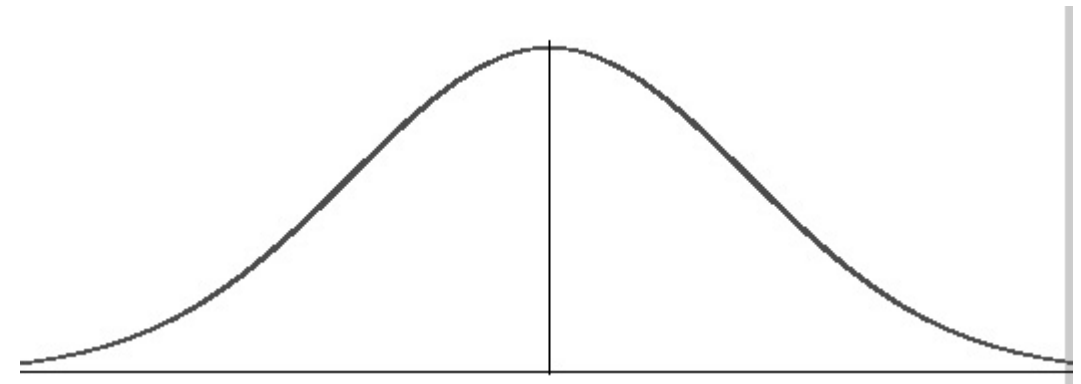


## Application 1 of multivariate calculus to probability/statistics: the Gaussian integral

The normal distribution (also called the Gaussian distribution) has probability density function proportional to  $e^{-x^2}$ , i.e.  $p(x) = \frac{1}{Z}e^{-x^2}$  for some  $Z$ . We need to choose the constant of proportionality  $Z$  so that the total probability is 1, i.e.

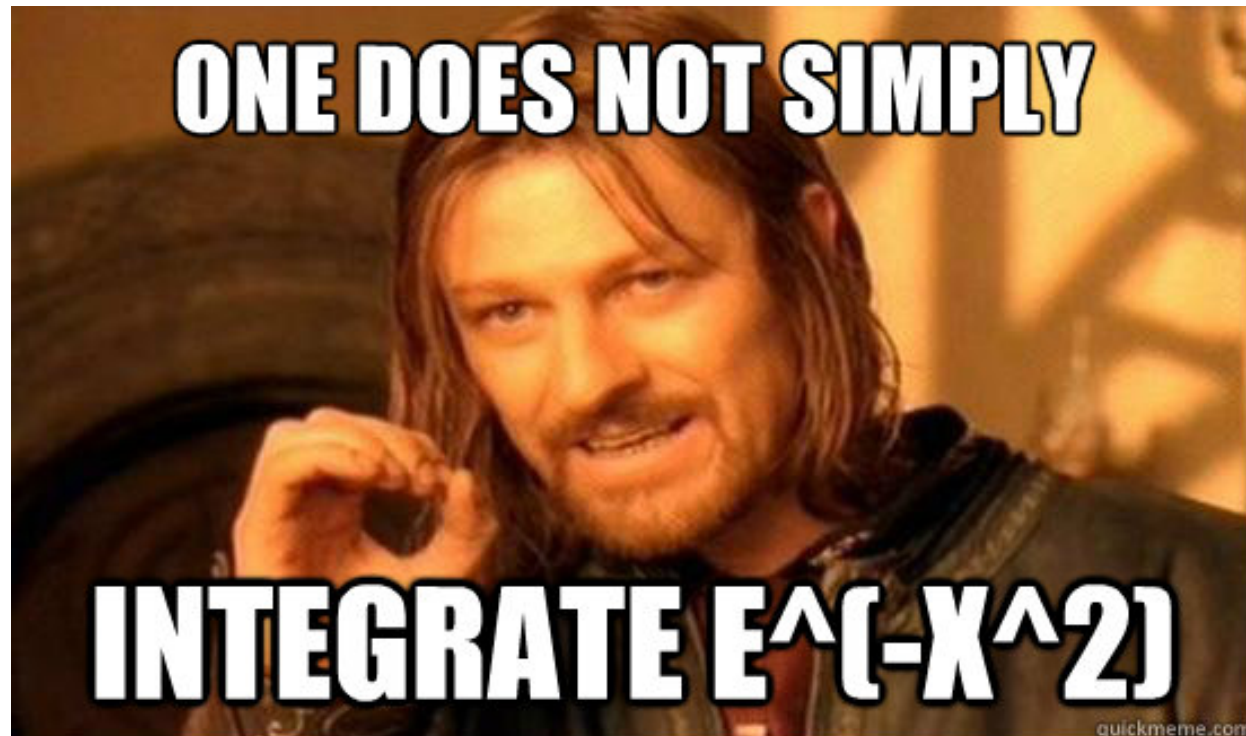
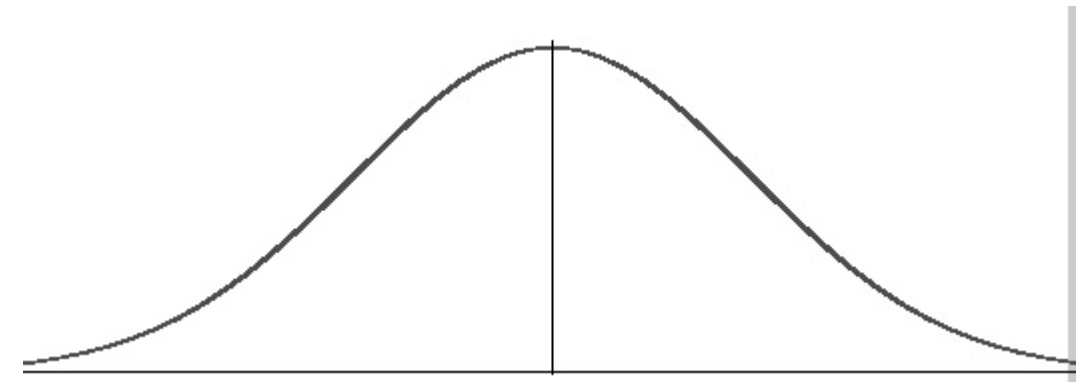
$$Z = \int_{-\infty}^{\infty} e^{-x^2} dx.$$



## Application 1 of multivariate calculus to probability/statistics: the Gaussian integral

The normal distribution (also called the Gaussian distribution) has probability density function proportional to  $e^{-x^2}$ , i.e.  $p(x) = \frac{1}{Z}e^{-x^2}$  for some  $Z$ . We need to choose the constant of proportionality  $Z$  so that the total probability is 1, i.e.

$$Z = \int_{-\infty}^{\infty} e^{-x^2} dx.$$



The problem is, we cannot write down the antiderivative of  $e^{-x^2}$  - it is not an elementary function. But multiple integration will help us in a surprising, clever way.

First, we need to show that the improper integral  $Z = \int_{-\infty}^{\infty} e^{-x^2} dx$  converges.

It will be enough to show that  $Z' = \int_1^{\infty} e^{-x^2} dx$  converges, because then

$$Z = \int_{-\infty}^{-1} e^{-x^2} dx + \int_{-1}^1 e^{-x^2} dx + \int_1^{\infty} e^{-x^2} dx = Z' + \int_{-1}^1 e^{-x^2} dx + Z'.$$

First, we need to show that the improper integral  $Z = \int_{-\infty}^{\infty} e^{-x^2} dx$  converges.

It will be enough to show that  $Z' = \int_1^{\infty} e^{-x^2} dx$  converges, because then

$$Z = \int_{-\infty}^{-1} e^{-x^2} dx + \int_{-1}^1 e^{-x^2} dx + \int_1^{\infty} e^{-x^2} dx = Z' + \int_{-1}^1 e^{-x^2} dx + Z'.$$

We show that  $Z'$  converges using the (non-examinable) technique of integral estimation by inequalities:

$e^{-x^2}$  is a non-negative function, so FTC1 says that  $F(R) = \int_1^R e^{-x^2} dx$  is an

increasing function (in  $R$ ).

So, using some theorems from analysis, we know that, if there is a number  $M$  such that

$M \geq F(R) = \int_1^R e^{-x^2} dx$  for every  $R > 1$ , then  $\lim_{R \rightarrow \infty} F(R)$  exists, i.e.  $Z'$  converges.

First, we need to show that the improper integral  $Z = \int_{-\infty}^{\infty} e^{-x^2} dx$  converges.

It will be enough to show that  $Z' = \int_1^{\infty} e^{-x^2} dx$  converges, because then

$$Z = \int_{-\infty}^{-1} e^{-x^2} dx + \int_{-1}^1 e^{-x^2} dx + \int_1^{\infty} e^{-x^2} dx = Z' + \int_{-1}^1 e^{-x^2} dx + Z'.$$

We show that  $Z'$  converges using the (non-examinable) technique of integral estimation by inequalities:

$e^{-x^2}$  is a non-negative function, so FTC1 says that  $F(R) = \int_1^R e^{-x^2} dx$  is an

increasing function (in  $R$ ).

So, using some theorems from analysis, we know that, if there is a number  $M$  such that

$M \geq F(R) = \int_1^R e^{-x^2} dx$  for every  $R > 1$ , then  $\lim_{R \rightarrow \infty} F(R)$  exists, i.e.  $Z'$  converges.

For all  $x \geq 1$ , we have  $e^{-x^2} \leq e^{-x}$ , so  $\int_1^R e^{-x^2} dx \leq \int_1^R e^{-x} dx = e^{-1} - e^{-R} \leq e^{-1}$ ,

so  $e^{-1}$  is the upper bound  $M$  that we want.

Reminder: we wish to evaluate  $Z = \int_{-\infty}^{\infty} e^{-x^2} dx$ .

Consider the double integral  $\iint_{\mathbb{R}^2} e^{-x^2} e^{-y^2} dA$ . The integrand is always positive, so we can calculate this improper integral using an iterated integral:

$$\iint_{\mathbb{R}^2} e^{-x^2} e^{-y^2} dA = \int_{-\infty}^{\infty} \int_{-\infty}^{\infty} e^{-x^2} e^{-y^2} dy dx = \int_{-\infty}^{\infty} e^{-x^2} Z dx = Z^2$$

Reminder: we wish to evaluate  $Z = \int_{-\infty}^{\infty} e^{-x^2} dx$ .

Consider the double integral  $\iint_{\mathbb{R}^2} e^{-x^2} e^{-y^2} dA$ . The integrand is always positive, so we can calculate this improper integral using an iterated integral:

$$\iint_{\mathbb{R}^2} e^{-x^2} e^{-y^2} dA = \int_{-\infty}^{\infty} \int_{-\infty}^{\infty} e^{-x^2} e^{-y^2} dy dx = \int_{-\infty}^{\infty} e^{-x^2} Z dx = Z^2$$

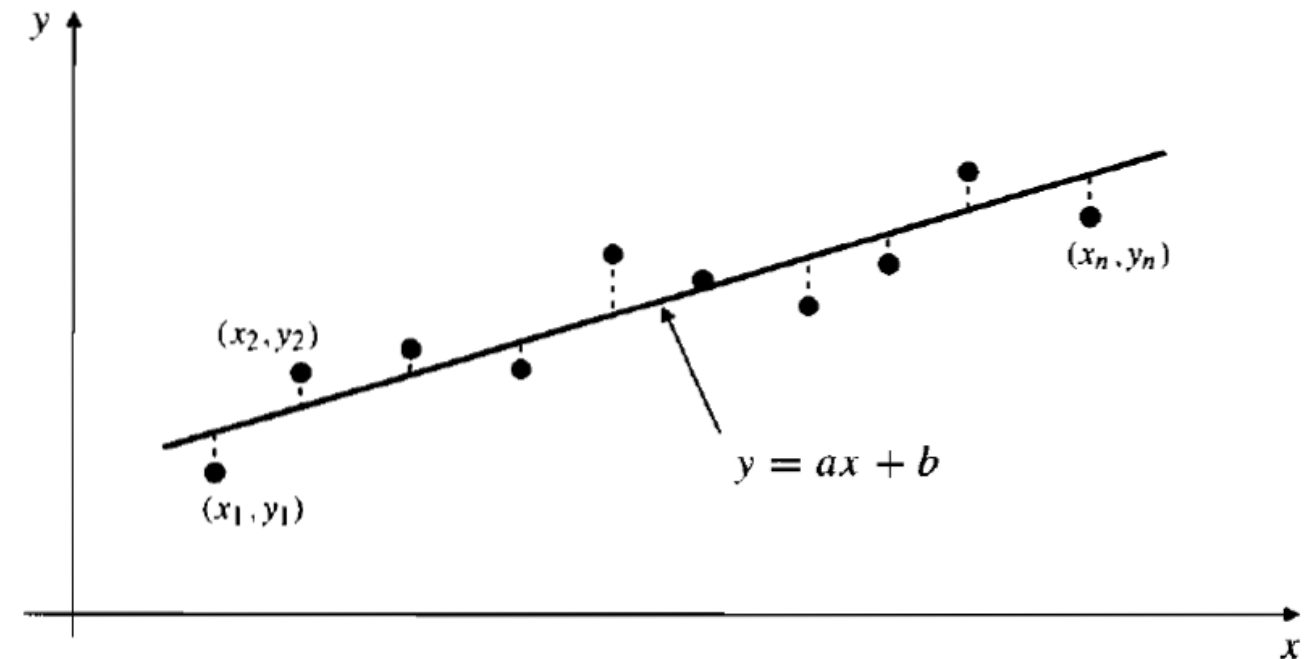
But we can also calculate this double integral using polar coordinates (yes, you can use change of variables on improper integrals):

$$\begin{aligned} \iint_{\mathbb{R}^2} e^{-x^2} e^{-y^2} dA &= \int_0^{2\pi} \int_0^{\infty} e^{-r^2} r dr d\theta \quad \text{inner integral independent of } \theta, \\ &\quad \text{and substitution } u = r^2 \\ &= 2\pi \int_0^{\infty} \frac{e^{-u}}{2} du = \pi \lim_{R \rightarrow \infty} \int_0^R e^{-u} du = \pi \left( \lim_{R \rightarrow \infty} 1 - e^{-R} \right) = \pi. \end{aligned}$$

So  $Z^2 = \pi$ , i.e.  $Z = \sqrt{\pi}$ .

## Application 2 of multivariate calculus to probability/statistics: linear regression

Suppose two physical quantities  $x$  and  $y$  (e.g. temperature and pressure) are related by  $y = ax + b$ , for some unknown constants  $a$  and  $b$ . To estimate  $a$  and  $b$ , we can do an experiment to find some data  $(x_1, y_1), \dots, (x_n, y_n)$ , then plot the data and draw a line that “best” fits the data. Mathematically, this means we want to maximise (or minimise) some function  $f(a, b)$ , where  $f$  measures “how well (or how badly) the line fits the data” - what this means will depend on the physical situation.





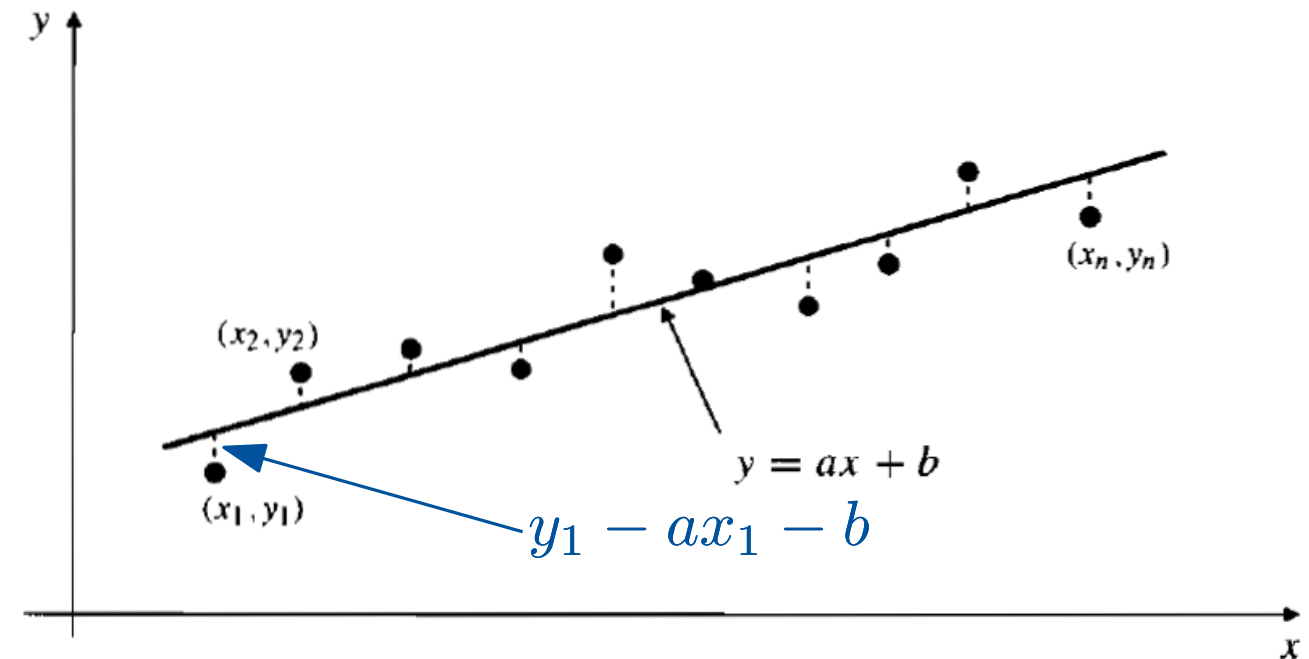
## Application 2 of multivariate calculus to probability/statistics: linear regression

Suppose two physical quantities  $x$  and  $y$  (e.g. temperature and pressure) are related by  $y = ax + b$ , for some unknown constants  $a$  and  $b$ . To estimate  $a$  and  $b$ , we can do an experiment to find some data  $(x_1, y_1), \dots, (x_n, y_n)$ , then plot the data and draw a line that “best” fits the data. Mathematically, this means we want to maximise (or minimise) some function  $f(a, b)$ , where  $f$  measures “how well (or how badly) the line fits the data” - what this means will depend on the physical situation.

One common and convenient scheme is **least squares**, where we minimise the error function

$$f(a, b) = \sum_{i=1}^n (y_i - ax_i - b)^2$$

the sum of squares of the vertical distances from the data points to the line (dotted lines in the diagram).



Reminder: given data points  $(x_1, y_1), \dots, (x_n, y_n)$ , we wish to minimise the error function  $f(a, b) = \sum_{i=1}^n (y_i - ax_i - b)^2$ .

(The notation is confusing:  $a, b$  are the unknowns, and  $x_i, y_i$  are known numbers.)

The domain for  $(a, b)$  is all of  $\mathbb{R}^2$ , which has no boundary. So, if a minimum for  $f$  exists, it must be at a critical point (because  $f$  is differentiable everywhere, so it has no singular points).

At a critical point:

$$\frac{\partial f}{\partial a} = 0 \Rightarrow \sum_{i=1}^n 2(y_i - ax_i - b)(-x_i) = 0 \Rightarrow \left( \sum_{i=1}^n x_i^2 \right) a + \left( \sum_{i=1}^n x_i \right) b = \sum_{i=1}^n x_i y_i$$

$$\frac{\partial f}{\partial b} = 0 \Rightarrow \sum_{i=1}^n 2(y_i - ax_i - b)(-1) = 0 \Rightarrow \left( \sum_{i=1}^n x_i \right) a + \left( \sum_{i=1}^n 1 \right) b = \sum_{i=1}^n y_i$$

Divide each equation on the far right hand side by  $n$ , and use the mean value notation  $\bar{x} = \frac{1}{n} \sum_{i=1}^n x_i$ :

$$\overline{x^2}a + \bar{x}b = \overline{xy}$$

$$\bar{x}a + b = \bar{y}$$

Combine into a matrix equation:

$$\begin{pmatrix} \overline{x^2} & \bar{x} \\ \bar{x} & 1 \end{pmatrix} \begin{pmatrix} a \\ b \end{pmatrix} = \begin{pmatrix} \overline{xy} \\ \bar{y} \end{pmatrix}$$

$$\text{So } \begin{pmatrix} a \\ b \end{pmatrix} = \begin{pmatrix} \overline{x^2} & \bar{x} \\ \bar{x} & 1 \end{pmatrix}^{-1} \begin{pmatrix} \overline{xy} \\ \bar{y} \end{pmatrix}, \text{ which means } a = \frac{\overline{xy} - \bar{x}\bar{y}}{\overline{x^2} - (\bar{x})^2}; \quad b = \frac{\overline{x^2}\bar{y} - \bar{x}\overline{xy}}{\overline{x^2} - (\bar{x})^2}.$$

Divide each equation on the far right hand side by  $n$ , and use the mean value notation  $\bar{x} = \frac{1}{n} \sum_{i=1}^n x_i$ :

$$\overline{x^2}a + \bar{x}b = \overline{xy}$$

$$\bar{x}a + b = \bar{y}$$

Combine into a matrix equation:

$$\begin{pmatrix} \overline{x^2} & \bar{x} \\ \bar{x} & 1 \end{pmatrix} \begin{pmatrix} a \\ b \end{pmatrix} = \begin{pmatrix} \overline{xy} \\ \bar{y} \end{pmatrix}$$

$$\text{So } \begin{pmatrix} a \\ b \end{pmatrix} = \begin{pmatrix} \overline{x^2} & \bar{x} \\ \bar{x} & 1 \end{pmatrix}^{-1} \begin{pmatrix} \overline{xy} \\ \bar{y} \end{pmatrix}, \text{ which means } a = \frac{\overline{xy} - \bar{x}\bar{y}}{\overline{x^2} - (\bar{x})^2}; \quad b = \frac{\overline{x^2}\bar{y} - \bar{x}\overline{xy}}{\overline{x^2} - (\bar{x})^2}.$$

To conclude that these values of  $(a, b)$  really minimises  $f(a, b) = \sum_{i=1}^n (y_i - ax_i - b)^2$ ,

we need to show that  $f$  achieves a minimum on  $\mathbb{R}^2$ . A precise proof is complicated.

The main idea: as  $(a, b)$  “moves away from the origin”,  $f(a, b) \rightarrow \infty$  (and  $\mathbb{R}^2$  has no boundary so we do not need to consider  $(a, b)$  moving towards a boundary point that is not in the domain).