Gaussian distribution

Gaussian distribution, also called "normal distribution," is the single most frequently used probability density function. It is more widely used in social sciences or experimental sciences than in theoretical physics. The Gaussian distribution is given as follows:

$$f(x, \mu, \sigma) = \frac{1}{\sqrt{2\pi}\sigma} e^{-\frac{(x-\mu)^2}{2\sigma^2}} \tag{1}$$

where μ is the mean (i.e. average) and σ is the standard deviation. See Fig. 1 for the graph of this probability density function. The peak is where the average is. This is the point where the chance is most likely. Notice also that the more you deviate from this peak the less probability density. You also see that the probability that the value falls between μ and $\mu + \sigma$ is 0.3413, between $\mu + \sigma$ and $\mu + 2\sigma$ is 0.1359, between $\mu + 2\sigma$ and $\mu + 3\sigma$ is 0.0214, bigger than $\mu + 3\sigma$ is 0.0013 and so on. Notice that the probability density rapidly decreases as the value deviates more from the mean. This is because of the presence of the exponential function which rapidly decreases as you plug in smaller and smaller negative numbers.

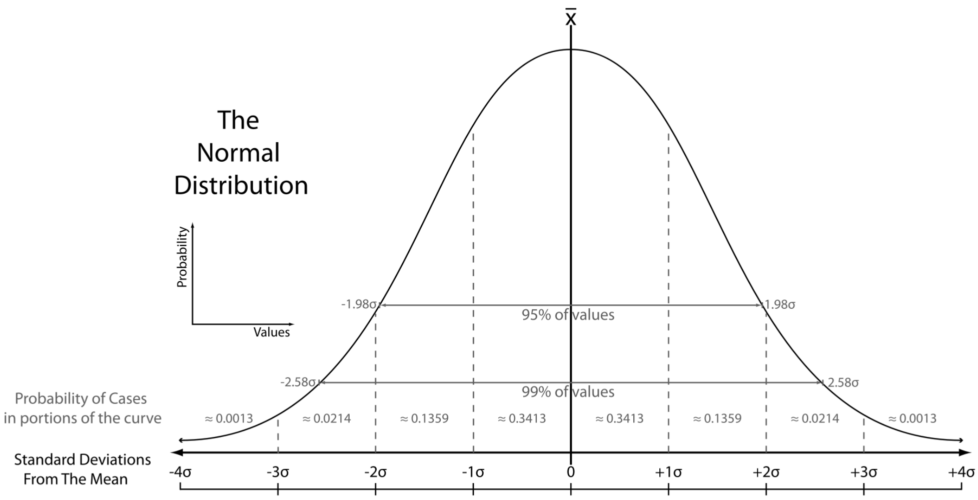

Figure 1: Gaussian distribution

Also, from the figure, we can see that the probability that you will get a value between $\mu - 1.98\sigma$ and $\mu + 1.98\sigma$ is 95%, while the probability that you will get a value between $\mu - 2.58\sigma$ and $\mu + 2.58\sigma$ is 99%. If you heard of "confidence interval" in a poll, it's exactly this. For example, let's say that there is a referendum and the bill will be passed if more than half of the electorate votes for yes. And, let's say that a poll was conducted, and the result was 48% yes with 95% confidence interval being between 46% and 50%. Then, the chance that the result will turn out to be less than 46% or more than 50% would be 5%. Since the chance that the result will be less than 46% is equal to the chance that the result will be less than 50% as the Gaussian distribution is symmetric, the former and the latter will be 2.5% each. Therefore, there is only 2.5% chance that the bill will be passed.

There is an important theorem related to Gaussian distribution called "central limit theorem." It states that under suitable (fairly common) condition the sum of many random values follow approximately Gaussian distribution. A good example is the binomial distribution, which is used for the poll as the one we just mentioned.

Let us give you another example. Let's say you cast an ordinary six sided die 10,000 times, and sum all the numbers you get. (We choose a rather big number 10,000, because due to the central limit theorem, the bigger this number, the closer the distribution to Gaussian distribution.) Then, the distribution of the sum you will get is roughly Gaussian distribution. Of course the mean of the sum is 3.5×10,000=35,000. (If you cast a die once, the expectation value is 3.5 as 1+2+3+4+5+6)/6=3.5.)

Problem 1. By consulting "Standard deviation of the sample means," check that the standard deviation of the sum is about 171.

Thus, the chance that the sum you obtain so is between $35,000-171$ and $35,000+171$ is about $0.3413 + 0.3413 = 0.6826$ (approximately 68%), as you can see from Fig. 1. In other words, if you divide this sum by 10,000, the number of toss, to obtain the average value, the chance that the average is between $3.5000 - 0.0171$ and $3.5000 + 0.0171$ is 68%.

Problem 2. What is the chance that the average so obtained, after 10,000 toss is bigger than $3.5513 (= 3.5000 + 0.0171 \times 3)$?

If you correctly solve this problem, you will get a very small number. Let's call it p. Then, if you cast a die 10,000 times and the average you get is bigger than 3.5513, there is only chance of p that the die is a good one, having the expectation value of 3.5.

Problem 3. Suppose you toss a coin 10,000 times. Assuming that the coin is an honest one, having the equal probability for head and for tail, what is the probability that you will get less than 4,900 heads? What is the probability that you will get more than 5,150 heads? Use the formulas in our earlier article "Binomial distribution."

Given all these, you are now able to understand if somebody says the discovery of

Higgs particle is confirmed at 5 sigma level. Even though we cannot see from the figure, if you look up a table for Gaussian distribution or use scientific calculator instead, the chance that something is randomly out of the range between $\mu - 5\sigma$ and $\mu + 5\sigma$ is 0.00000057 (or 1 in 1.7 million) for Gaussian distribution. This implies that the chance that the Higgs signal we found was due to just background random event is about 1 in 1.7 million.

In April 2021, it was announced that there is a 4.2 sigma difference between the experimental value and the theoretical value of the anomalous magnetic moment of muon. (If you want to know what is the anomalous magnetic moment, please read "Electron magnetic moment.") Let's see how this number 4.2 is calculated. The theoretical value was 0.001 165 918 10(43) and the experimental value was 0.001 165 920 61(41) [1, 2].[1] So, you may think that the theory and the experiment agree very well. But, no. The experimental value is bigger than the theoretical value by $(251 \pm 59) \times 10^{-11}$ $(92061 - 91810 = 251)$ and $\sqrt{43^2 + 41^2} \approx 59$. So, if you divide 251 by 59, you get 4.2. Actually, there is only 1 in 180 million chance that a value in a Gaussian distribution lies outside 4.2 standard deviation range. The theoretical calculation was based on the Standard Model. Therefore, assuming that the theoretical calculation was performed by properly following the Standard Model, there is only 1 in 180 million chance that the Standard Model is correct. Therefore, we can safely say that either the Standard Model is wrong, or the theoretical calculation didn't properly follow the Standard Model. (There was a paper that claimed that the latter was the case.)

Problem 4. Check that the distribution (1) yields μ as mean and σ as standard deviation.

Summary

- Gaussian distribution (also called "normal distribution") is given in the form of

$$e^{-\frac{(x-\mu)^2}{2\sigma^2}}$$

  where μ is the mean and σ is the standard deviation.

- Central limit theorem says that, if you repeat an experiment many times, add the obtained results, their distribution approaches the Gaussian distribution.

References

[1] T. Aoyama, N. Asmussen, M. Benayoun, J. Bijnens, T. Blum, M. Bruno, I. Caprini, C. M. Carloni Calame, M. Cè and G. Colangelo, *et al.* "The anomalous mag-

[1] If you are not familar with this kind of notation, please read "Scientific notation."

netic moment of the muon in the Standard Model," Phys. Rept. **887**, 1-166 (2020) doi:10.1016/j.physrep.2020.07.006 [arXiv:2006.04822 [hep-ph]].

[2] B. Abi *et al.* [Muon g-2], "Measurement of the Positive Muon Anomalous Magnetic Moment to 0.46 ppm," Phys. Rev. Lett. **126**, no.14, 141801 (2021) doi:10.1103/PhysRevLett.126.141801 [arXiv:2104.03281 [hep-ex]].