1 Exponential function

The *exponential function* exp is defined by the power series

$$\exp(x) = \sum_{k=0}^{\infty} \frac{1}{k!} x^k = 1 + x + \frac{x^2}{2} + \frac{x^3}{6} + \frac{x^4}{24} + \cdots .$$

This function enjoys several remarkable properties.

- It relates addition to multiplication: $\exp(x + y) = \exp(x) \cdot \exp(y)$.
- It's its own derivative: $\frac{d}{dx} \exp(x) = \exp(x)$. Then of course $\int \exp(x)\, dx = \exp(x) + C$.
- If you plug imaginary numbers into exp, then trigonometry emerges:

  $$\exp(i\theta) = \cos(\theta) + i \sin(\theta).$$

  That's surprising, because the power series above does not apparently have anything to do with trigonometry. (In fact, this is just a hint at a deeper phenomenon in higher mathematics: The exponential function maps Lie algebras to their associated Lie groups.)

Exercise 1. *Show that* exp *is its own derivative.*

There is a second, equivalent way to define the exponential function:

$$\exp(x) = \lim_{n\to\infty} \left(1 + \frac{x}{n}\right)^n .$$

This definition turns out to be useful in some problems of computing limits.

The inverse function of exp is denoted log. It also enjoys useful properties, including $\log(x \cdot y) = \log(x) + \log(y)$. The addition-multiplication behavior of exp and log is the basis for many computations throughout mathematics, statistics, and computer science.

2 Interlude: e

Whether you know it or not, your favorite number is 0. When you plug 0 into exp, you get $\exp(0) = 1$, which is your second-favorite number. So what happens when you plug your second-favorite number into exp? You get a weird number $\exp(1) = 2.718\ldots$. Just by virtue of being $\exp(1)$, this number deserves a name. We call it e.

Using the addition-multiplication relationship above, one can prove that $\exp(n) = e^n$ for all integers n. Based on this fact, we *define* e^x, for all real numbers x, to be $\exp(x)$. But let me emphasize that you should view exp as the primary object, and e as a secondary phenomenon that emerges from exp. Similarly, one can think of log as a logarithm with base e, but it is ultimately more useful to think of it as the inverse to the exponential function.

3 Poisson distribution

A discrete random variable X is *Poisson-distributed* with parameter λ, written $X \sim \text{Pois}(\lambda)$, if

$$P(X = k) = \frac{e^{-\lambda}\lambda^k}{k!}$$

for $k = 0, 1, 2, \ldots$ (and $P(X = k) = 0$ for other k).

Exercise 2. *Show that the probability mass function sums to* 1*.*

Exercise 3. *Compute the expectation* $E(X)$*.*

Exercise 4. *Compute the variance* $V(X)$*. (Differentiation is required.)*

4 Applications

The Poisson distribution is useful for modeling a specific, extremely common situation. Suppose that a certain kind of event occurs uniformly randomly with respect to time. Suppose that the number of occurrences doesn't depend on anything but the length of the time interval of observation. Suppose further that it is practically impossible for multiple occurrences to happen at exactly the same time. Then the number of occurrences in a fixed time period can be modeled as a Poisson-distributed X. Usually, we have some information about the average number of occurrences per time period, and this information helps us determine λ.

For example, we might be talking about lightning strikes that will occur this year. Assume that the number of strikes this year doesn't depend on the number of strikes last year, whether we measure years from January 1 to December 31 or from July 1 to June 30, etc. Assume that no two strikes are exactly simultaneous. Then the number of strikes is Poisson-distributed, and meteorological records let us estimate λ.

Exercise 5. *Over the past* 43 *years, the USA has experienced* 33 *"major" (meaning magnitude 7 or greater, I think) earthquakes. If we model the number of earthquakes next year as a Poisson random variable* X*, then what is* λ*? What is the meaning and the numerical value of* $P(X \geq 2)$*?*

The 1990s book and film *A Civil Action* tell a true story about the town of Woburn, Massachusetts. Residents of the town sue a local company because they suspect that the company's pollution is causing elevated cancer rates in the town.

Exercise 6. *Suppose that the USA has about* 280,000,000 *people and* 30,800 *leukemia cases annually. So there are* 30,800/280,000,000 $\approx$ 0.00011 *cases per person. Now consider a town of population* 35,000*. Assuming that leukemia rates are the same in this town as in the USA, how many cases does one expect in the town this year? What is the probability that the town has* 8 *or more cases this year?*

5 Relationship to the binomial distribution

A binomially distributed random variable Y counts the number of successes in n independent Bernoulli trials, each of success probability p. One might ask, "How many successes will occur in infinitely many trials?" Intuitively, the answer should be something like $\infty \cdot p = \infty$. So it's not a good question.

However, there is a way to tweak the question so that it gives a sensible, finite answer. As we let the number n of trials go to infinity, we balance it by letting the success probability p go to zero. More precisely, fix a number λ and set $p = \lambda/n$ in the binomial:

$$\begin{aligned} P(Y = k) &= \binom{n}{k} p^k (1-p)^{n-k} \\ &= \frac{n!}{k!(n-k)!} \left(\frac{\lambda}{n}\right)^k \left(1 - \frac{\lambda}{n}\right)^{n-k} \\ &= \frac{\lambda^k}{k!} \frac{n(n-1)\cdots(n-k+1)}{n^k} \left(1 - \frac{\lambda}{n}\right)^n \left(1 - \frac{\lambda}{n}\right)^{-k}. \end{aligned}$$

Exercise 7. *As $n \to \infty$, what happens to each of the four factors in the expression above?*

The preceding exercise should show that the binomial limits to the Poisson. Practically speaking, when n is large, p is small, and $\lambda = pn$ is medium-sized, the Poisson distribution is a good approximation to the binomial distribution. In other words, the Poisson is useful for modeling rare events.

Exercise 8. *Suppose that $X_1 \sim \text{Pois}(\lambda_1)$ and $X_2 \sim \text{Pois}(\lambda_2)$ are independent. Find the probability mass function of $X_1 + X_2$.*