Large Deviations and Rare Events: A Very Short Summary[1]

The setting. Given a particular model, we try to answer the following questions:

(1) What is a typical behavior of the system?
(2) What is behavior is non-typical and how unlikely is it?
(3) If the system does exhibit non-typical behavior, what is the most likely scenario?

An example. The system is $\{\sqrt{\varepsilon}\, W(t),\ t \in [0,1]\}$, where W is a standard Brownian motion and ε is a small parameter. Then

(1) The typical behavior of the system is to stay near zero:
$$\lim_{\varepsilon\to 0} \mathbb{P}\Big(\sqrt{\varepsilon} \sup_{0<t<1} |W(t)| = 0\Big) = 1$$
(2) Non-typical behavior is to drift away from zero, and it is exponentially unlikely: for a function $\varphi : [0,1] \to \mathbb{R}$,
$$\lim_{\delta\to 0}\lim_{\varepsilon\to 0}\ \varepsilon\ \ln \mathbb{P}\Big(\sup_{0<t<1} |\sqrt{\varepsilon} W(t) - \varphi(t)| < \delta\Big) = -I(\varphi),$$
where
$$I(\varphi) = \begin{cases} \frac{1}{2}\int_0^1 |\varphi'(t)|^2\, dt, & \text{if } \varphi(0)=0 \text{ and } \varphi' \text{ exists} \\ +\infty, & \text{otherwise} \end{cases}$$
(3) If, for small ε, the trajectory of $\sqrt{\varepsilon}\, W$ gets near a point $a \neq 0$ at time $t = 1$, then this trajectory will most likely follow the line $y = at$. The reason is that the linear function from 0 to a minimizes $I(\varphi)$ over all absolutely continuous functions on $[0,1]$ that start at 0 and end at a.

This example is the consequence of the Donsker theorem establishing the large deviations principle (LDP) for the Brownian motion.

Of course, in the case of the Brownian motion, we know the distribution of the hitting time
$$T_\varepsilon = \inf\{t > 0 : \sqrt{\varepsilon} W(t) = 1\};$$
in particular,
$$\mathbb{P}(T_\varepsilon < t) = \frac{2}{\sqrt{2\pi}} \int_{1/\sqrt{\varepsilon t}}^{\infty} \exp\left(-\frac{x^2}{2}\right) dx,$$
which, in particular, shows that even though $\mathbb{P}(T_\varepsilon < \infty) = 1$, we have
$$\mathbb{P}(T_\varepsilon \sim 1) \sim e^{-1/(2\varepsilon)},\ \ \mathbb{P}(T_\varepsilon \sim 1/\sqrt{\varepsilon}) \sim 1.$$
The theory of large deviations makes similar conclusions possible even when no closed-form expressions are available.

For a concrete numerical example, take $\varepsilon = 10^{-4}$ so that $\sqrt{\varepsilon} = 1/100$ and $1/(2\varepsilon) = 5000$. Then we know for sure that a trajectory of $W(t)$ starting at zero will eventually reach the level $\sqrt{1/\varepsilon} = 100$,

[1] Sergey Lototsky, USC; version of May 13, 2024

but it will typically take about 100 units of time; the probability that it happens in unit time is very small: of order e^{-5000}.

A road map.

(1) Set up the problem so that typical behavior $\bar{X}$ is non-random and corresponds to a law of large numbers, with the corresponding asymptotic parameter ε. Part of the setting is the metric $\rho = \rho(X_\varepsilon, Y)$ that quantifies how close the system X_ε is to a particular realization Y.
(2) Get a result of the form

$$\lim_{\delta\to 0}\lim_{\varepsilon\to 0} \varepsilon \ \ln \mathbb{P}\big(\rho(X_\varepsilon, Y) < \delta\big) = -I(Y), \ I(Y) \geq 0, \ I(\bar{X}) = 0.$$

Together with a technical condition known as *exponential tightness*, this will imply that large deviations principle holds.
(3) Given a characterizing property $\mathcal{P}$ of a non-typical event (like reaching the point a at time 1 in the Brownian motion example), we then conclude that, in the unlikely case that $\mathcal{P}$ happened, the system will follow "the most likely of all the unlikely realizations"

$$\tilde{Y} = \text{argmin}\{I(Y) : Y \in \mathcal{P}\}.$$

A challenge: what if "typical behavior" is a set rather than one particular outcome? [The good news is that there are some results about LDP for set-valued systems.]