Some facts about Weak Convergence :

Suppose E is a Polish Space. & $M_1(E)$ is the space of probability measures on E. We have a topology on $M_1(E)$, which although is called "weak", ~~that~~ is rather rich and yields a nice metric, which we can work with and makes $M_1(E)$ a complete, separable metric space. ~~Then one can~~ ~~prove~~ Let us say that $(\mu_n)_{n\in\mathbb{N}}$ converges weakly to μ in $M_1(E)$ iff $\forall f \in C_b(E)$,

$$\int f d\mu_n \to \int f d\mu.$$

This notion of convergence is compatible with respect to the topology generated by sets (basis of the topogy)

$$U_f(x,\delta) = \left\{ \nu \in M_1(E) : \left| \int_E f d\nu - x \right| < \delta \right\} \text{ for } f \in C_b(E).$$

$\delta > 0$ & $x \in \mathbb{R}$.

<u>Thm</u> : If E is Polish, then so is $M_1(E)$ under the above topology. Furthermore, $M_1(E)$ is compact if and only if E is compact.

The following theorem is at the heart of many limit theorems in probability theory.

<u>Thm (Prohorov)</u>: A family $\{\mu_\varepsilon\}_{\varepsilon > 0}$ of prob. measures of $M_1(E)$ is compact if and only if it is uniformly tight.

Here, recall that compactness requires, given any sequence there is some subsequence which has a weak limit. ~~[illegible]~~

Also recall that ~~unifor~~ $\{\mu_\varepsilon\}_{\varepsilon>0}$ uniformly tight means: ~~there must be a compact set~~ ~~$K \subset E$ (K must be indepen~~ given any $\eta > 0$ arbitrarily small, there must be compact set $K = K(\eta)$ (independent of $\varepsilon > 0$, uniformity!) so that $\mu_\varepsilon(K_\eta^c) \leq \eta \ \forall \varepsilon > 0$.

Do some examples:

Determine if the following sequences in $M_1(\mathbb{R})$ converge weakly or not:

1) $\mu_n = \delta_{1/n}$
2) $\mu_n = (1-\frac{1}{n})\delta_n + \frac{1}{n}\delta_0$
3) $\mu_n = \delta_n$
4) $\mu_n =$ Gaussian with 0 mean & variance n.
5) $\mu_n =$ Gaussian with mean n & variance 1.
6) $\mu_n = \frac{1}{3}\underbrace{\mathcal{N}(\frac{1}{n},1)}_{\text{Gaussian with mean } \frac{1}{n}\text{, variance 1}} + \frac{1}{3}\underbrace{\mathcal{N}(n,1)}_{\text{Gaussian, mean } n\text{, Variance 1}} + \underbrace{\frac{1}{3}\mathcal{N}(0,1)}_{\text{Gaussian, mean 0, Variance 1}}$.
7) $\mu_n = \mathrm{Unif}([e^{-n}, e^n])$.

The following equivalent characterization is sometimes useful:

<u>Theorem</u> (<u>Portmanteau</u>): Let E be Polish. Then the following are equivalent.

(1) $\mu_n \Rightarrow \mu$ in $M_1(E)$.

(2) $\forall f \in C_b(E)$, $\lim_n \int f \, d\mu_n = \int f \, d\mu$

(3) $\forall f$ bounded & <u>uniformly continuous</u>, $\lim_n \int f d\mu_n = \int f d\mu$

(4) $\forall F \subset E$ closed, $\limsup_{n\to\infty} \mu_n(F) \leq \mu(F)$

(5) $\forall G \subset E$ open, $\liminf_{n\to\infty} \mu_n(G) \geq \mu(G)$

(6) For any set $A \subset E$ for which $\mu(\bar{A} \setminus A^\circ) = 0$ (such a set is called a μ-continuity set), we have $\lim_{n\to\infty} \mu_n(A) = \mu(A)$. $\boxed{/\!/}$

Now, look back at (4) & (5) (and also (6)) and recall the definition of a large deviation principle. May be this helps you to convince (at least partially) that a definition of a large deviation principle did not fall from the sky!