（a）

**Describe the curse of dimensionality**

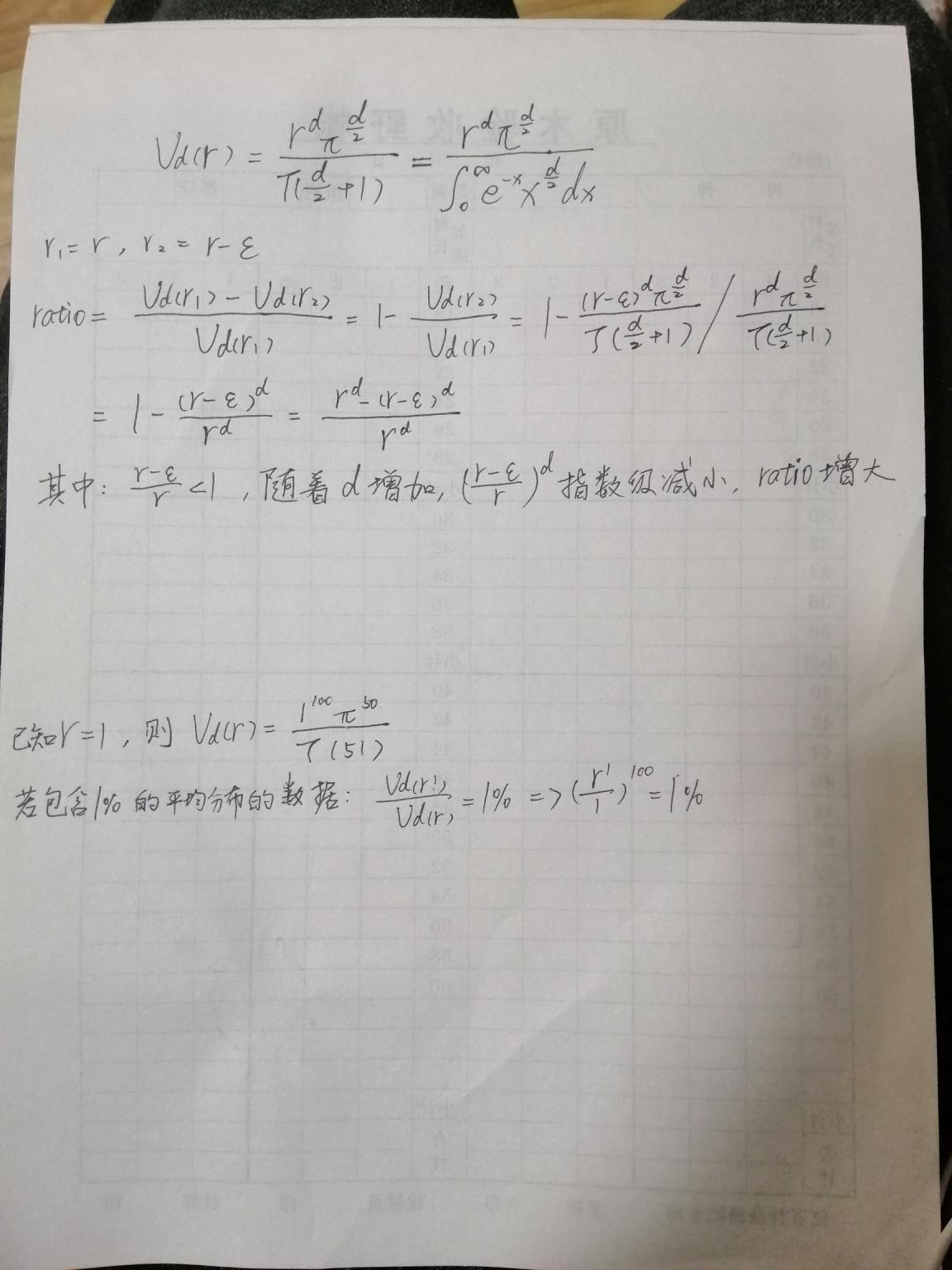
For the data set whose number of samples has been determined, there is a maximum value of the feature dimension that can make the best classification effect. When the number of the feature dimension exceeds this maximum value, the performance of the classifier is not improved, but degraded.

**Why does it make learning diffiffifficult in high dimensional spaces?**

The core reason is the mismatch between the original probability model and the actual probability model. Mathematically, if we want the probability model to match the actual situation as much as possible, with the increase of dimensions, the needed number of samples is exponential growth. However, the actual situation is that the number of samples is difficult to grow exponentially with the increase of dimensions, which leads to dimension disaster.

（b）

**What is the ratio between the volume of the crust and the volume of the hypersphere?**

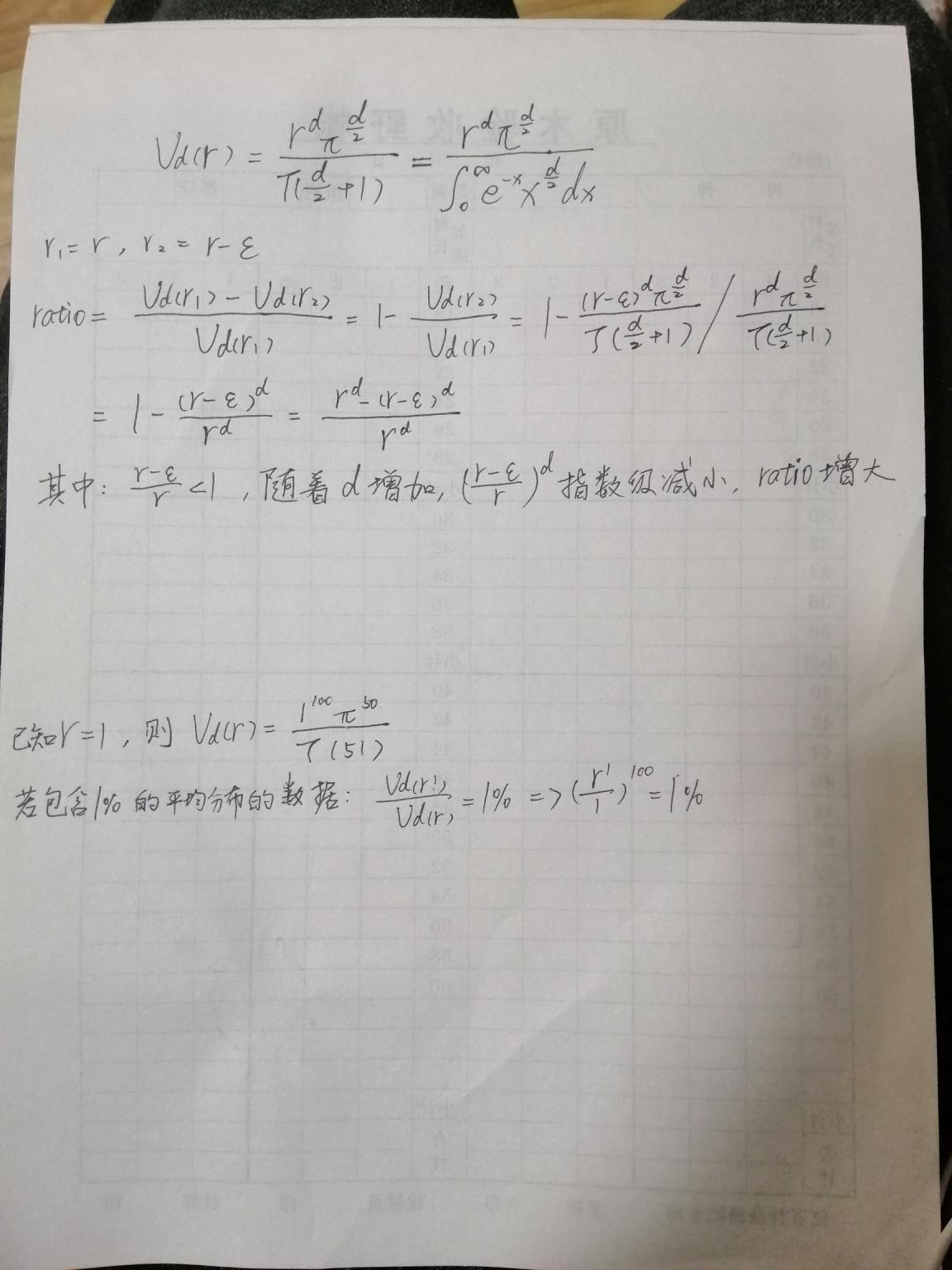


**How does the ratio change as *d* increases?**

, with the increase of d,  decreases exponentially and the ratio increases rapidly

(c)

**How big should *r0* be to ensure that the hypersperical neighborhood contains 1% of the data (on average)? How big to contain 10%?**



contains 1%:

contain 10%: