1. 泛化vs过拟合
   1. 泛化：
      1. 我们从分布中随机抽取****独立同分布**** (****i.i.d****) 的样本。【不能有相关因素】
      2. 分布是****平稳的。【不能受周期性因素影响】****
      3. 从**同一分布**的数据划分中抽取样本
   2. 过拟合：过拟合模型在训练过程中产生的损失很低，但在预测新数据方面的表现却非常糟糕。
2. 机器学习目标：从真实概率分布（已隐藏）中抽取的新数据做出良好预测。

奥卡姆剃刀定律：“如无必要，勿增实体”

1. 训练集和测试集：拆分数据。
   1. 原则1：测试集足够大
   2. 原则2：能代表整个数据集
   3. 勿对测试数据进行训练。
2. 测试速率、数据比率与结果的关系
   1. 速率过高会导致测试损失明显高于训练损失。
   2. 将训练数据与测试数据之比从 50% 降至 10% 大幅降低了训练集中数据点的个数。由于数据太少，较高的批量大小和学习速率会导致训练模型沿着曲线无规律地跳动（在最低点上方反复跳跃）。