**数据集说明**

**Labtest数据**

Labtest为一样例多label数据集。

打开数据集：f = h5py.File(‘Lab.h5’, ‘r’)

查看内容：f.keys()



如图可见，有五张表

Event：



长度为5000的事件序列，其余同death，略

Time：

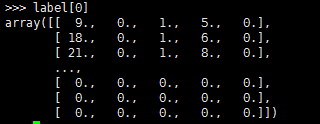


长度为5000的时间序列，其余同death，略

Label：



Label表的维度为[样例个数, 3682, 5]，每行对应一个label列表，表示该行的事件序列和时间序列唯一对应的label列表。长度为3682，每个label为一个五元组。



如上图，每个label序列由若干个label组成，每个label为一个五元组，分别表示label的labtest类型编号，labtest数值，labtest分类指标【0-normal，1-abnormal，2-delta】，labtest预测序列的结束位置，labtest预测序列的开始位置。

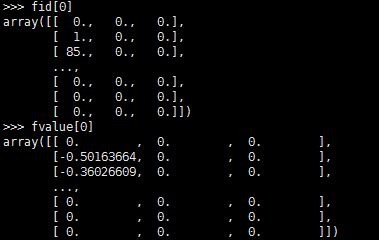
如第一行表示一个label，其对应的lab事件为9号事件，化验数值为0，验证类型为1-abnormal，这个label的预测对应着从0到5这段序列，也即我们以后将用事件和时间序列中0-5这部分之间的信息来预测这个label的相关信息。

Feature\_id和feature\_value：





Feature\_id和feature\_value分别与death数据集中feature表的1,3,5行和feature表中的2,4,6行对应，此处不再赘述。



如上图，第二个表示第二个事件有一个feature，id为1，数值为-0.50163664【已正规化】