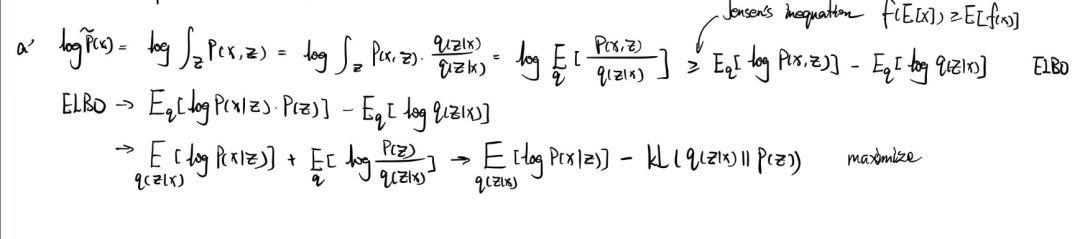
#### VAE的ELBO项推导（有多种推导方式，变分法，jensen不等式等）：



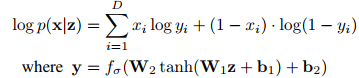
#### How to deal with the ELBO first item？(为什么以及什么情况用重构误差处理)

几个处理：1）Monte Carlo采样计算期望; 2）重参数化trick, z采样：

C:\Users\lenovo\AppData\Local\Temp\1582552563(1).png ≌ C:\Users\lenovo\AppData\Local\Temp\1582554157(1).png，经过重参数化trick，z采样过程中随机性与网络中参数无关，不会导致后向传播过程中梯度计算的variance.

目前，ELBO第一项转换为对C:\Users\lenovo\AppData\Local\Temp\1582554157(1).png的计算，其中：

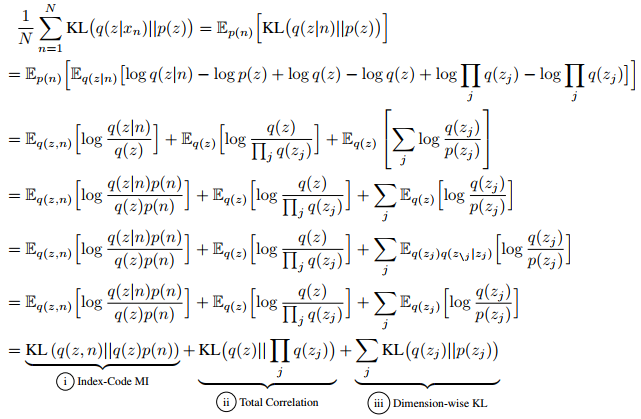
对的处理：

1. Gaussian:
2. Bernoulli: 

#### FactorVAE和β-TCVAE：对KL项的分解

KL项的分解：

β-TCVAE中KL项的分解：



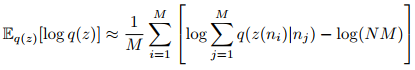
FactorVAE中目标函数：

β-TCVAE中目标函数与FactorVAE等价，即添加正则项；

##### 下一步，计算：

**β-TCVAE中估算：**

Minibatch Weighted Sampling方法：



其中， 是从分布中的采样; 推导过程借鉴了重要性采样的思想。把q(z)看成是q(z|n)的混合分布，其中n为混合index. 假如是z的重要来源成分，则应该有一个较大值;

但是详细推导过程没看明白；

**FactorVAE中计算：**

D(z): probability of sample z is from q(z) rather than .

推导过程用到了density ratio trick. 判别z属于q(z)还是，训练判别器,

对于最优判别器：，则对应的最优判别器，即得出F1式.

#### Z先验分布的选择及处理；

#### Metrics-1：FactorVAE score

#### Metrics-2：MIG (Mutual Information Gap)