問 2　生のデーテから求めた平均値$\overline{x}$と，度数分布表から求めた平均値$\overline{x}'$との差は，どのように考えればよいのですか。

(答え)　度数分布表の階級の幅をcとしますと，$|\overline{x}-\overline{x}'|\leqq\frac{c}{2}$となることは，容易に証明できます。つまり，次の不等式が成立します。

$$\overline{x}'-\frac{c}{2}\leqq\overline{x}\leqq\overline{x}'+\frac{c}{2}$$

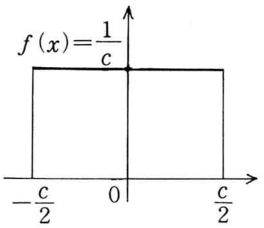
例えば，c＝10としますと，$\overline{x}$と$\overline{x}'$との間には，最大限±５の差があることになり，ちょっと，大きなずれのような気がします。したがって，できるだけくわしく平均値を求めたいときには，階級の幅cを小さくしなければなりません。しかし，逆にcを小さくしますと，度数分布表から，全体の傾向がつかみにくくなります。

要するに，そのデータから，どの程度のことが知りたいのか，そこをよく確かめてから度数分布表をつくるようにします。どうしてもくわしい平均値を知る必要があるときには，生のデータから平均値を計算しなければなりません。

ちょっと固いことをいいましたが，実際には，nがある程度大のとき，$\overline{x}$と$\overline{x}'$との差はそう大きなものにはならないようです。

また，この理論的背景には，次のことがあげられます。

階級内の各値の誤差Xiが一様分布をする，と仮定しますと，



◦ X_iの平均値$m=\int_{-\frac{c}{2}}^{\frac{c}{2}}xf(x)\,dx=\frac{1}{c}\int_{-\frac{c}{2}}^{\frac{c}{2}}x\,dx=0$

◦ X_iの分散$\sigma^2=\int_{-\frac{c}{2}}^{\frac{c}{2}}(x-m)^2f(x)\,dx$

$$=\frac{1}{c}\int_{-\frac{c}{2}}^{\frac{c}{2}}x^2\,dx=\frac{c^2}{12}$$

したがって，nが十分大のときには，誤差の平均値$\overline{X}$の分布は，正規分布N（0，$\frac{c^2}{12}$）に近似します。そして，誤差の平均値と，度数分布表から得られた平均値の誤差Eとは一致しますので，