

$$\text{信頼度95\%で, } |E| \leq 2\sqrt{\frac{c^2}{12}} = \frac{c}{\sqrt{3n}}$$

となります。

例えば,  $c=10$ ,  $n=100$ のとき,  $\bar{x}$ と $\bar{x}'$ との最大誤差は $\frac{10}{2}=5$ ではありませんが, 信頼度95%で,

$$|E| \leq \frac{c}{\sqrt{3n}} = \frac{10}{\sqrt{300}} \doteq 0.58$$

となります。

また, 階級の数 $k$ が12以下の場合で, データがほぼ正規分布をなすとき,

$$(\text{度数分布表から得られた分散 } s'^2) - \frac{c^2}{12}$$

を分散の値とした方が適当であるといわれており, これをシェパードの修正といいます。しかし, とくにくわしい分散(標準偏差)の値が必要な場合は, 生のデータから計算すればよいので, ふつうは, 度数分布表から得られたままの値で, 十分に合いますから, この修正は必要ありません。

下の表は, 実際の例で, 生のデータからと, これを度数分布表にまとめてから求めた平均値, 標準偏差などをまとめたものです。上で述べたことがらに, きわめてよく合っていることがわかります。

(表)

	例 1	例 2	例 3	例 4
データ数 $n$	479	496	518	496
満点   階級の幅 $c$	50   5	50   5	50   5	40   4
度数分布表からの平均値 $\bar{x}$	29.158	26.518	30.384	22.556
度数分布表からの標準偏差 $s'$	9.690	9.217	10.036	9.613
生のデータからの平均値 $\bar{x}$	29.207	26.516	30.485	22.486
生のデータからの標準偏差 $s$	9.614	8.997	9.928	9.519
平均値の差 $ \bar{x} - \bar{x}' $	0.049	0.002	0.101	0.070
$c / \sqrt{3n}$	0.132	0.130	0.127	0.104
$\sqrt{s'^2 - \frac{c^2}{12}}$	9.583	9.103	9.932	9.543