問4　二組のデータA_1，A_2 を合わせた場合の平均値と分散とは，どうなりますか。

(答え)　大きさn_1のデータA_1の平均値を$\overline{x_1}$，標準偏差をs_1

大きさn_2のデータA_2の平均値を$\overline{x_2}$，標準偏差をs_2

A_1，A_2を合わせた大きさ$n=n_1+n_2$の平均値を$\overline{x}$，標準偏差をs としますと，

$$\circ\ \overline{x}=\frac{n_1x_1+n_2x_2}{n_1+n_2}$$

$$\circ\ s^2=\frac{1}{n_1+n_2}\ \{\sum_{i=1}^{n_1}(x_i-\overline{x})^2+\sum_{j=1}^{n_2}(x_j-\overline{x})^2\}$$

$$=\frac{1}{n_1+n_2}\left[\sum_{i=1}^{n_1}\{(x_i-\overline{x_1})+(\overline{x_1}-\overline{x})\}^2+\sum_{j=1}^{n_2}\{(x_j-\overline{x_2})+(\overline{x_2}-\overline{x})\}^2\right]$$

$$=\frac{1}{n_1+n_2}\left[\sum_{i=1}^{n_1}(x_i-\overline{x_1})^2+2(\overline{x_1}-\overline{x})\sum_{i=1}^{n_1}(x_i-x_1)+\sum_{i=1}^{n_1}(\overline{x_1}-\overline{x})^2\right.$$

$$\left.+\sum_{j=1}^{n_2}(x_j-\overline{x_2})^2+2(\overline{x_2}-\overline{x})\sum_{j=1}^{n_2}(x_j-\overline{x_2})+\sum_{j=1}^{n_2}(\overline{x_2}-\overline{x})^2\right]$$

ここで　$\sum_{i=1}^{n_1}(x_i-\overline{x_1})=0$，　$\sum_{j=1}^{n_2}(x_j-\overline{x_2})=0$

$$s_1^2=\frac{1}{n_1}\sum_{i=1}^{n_1}(x_i-\overline{x_1})^2,\qquad s_2^2=\frac{1}{n_2}\sum_{j=1}^{n_2}(x_j-x_2)^2$$

$$\therefore s^2=\frac{1}{n_1+n_2}\left[n_1s_1^2+n_1(\overline{x_1}-\overline{x})^2+n_2s_2^2+n_2(\overline{x_2}-\overline{x})^2\right]$$

上の式もよく用いられますが，更にこれを変形して，

$$n_1(\overline{x_1}-\overline{x})^2=n_1(\overline{x_1}-\frac{n_1\overline{x_1}+n_2\overline{x_2}}{n_1+n_2})^2=\frac{n_1n_2^2(\overline{x_1}-\overline{x_2})^2}{(n_1+n_2)^2}$$

$$n_2(\overline{x_2}-\overline{x})^2=n_2(\overline{x_2}-\frac{n_1\overline{x_1}+n_2\overline{x_2}}{n_1+n_2})^2=\frac{n_1^2n_2(\overline{x_1}-\overline{x_2})^2}{(n_1+n_2)^2}$$

$$\therefore s^2=\frac{n_1s_1^2+n_2s_2^2}{n_1+n_2}+\frac{n_1n_2(\overline{x_1}-\overline{x_2})^2}{(n_1+n_2)^2}$$

この式も，よく用いられます。