Question

Establish the following recursion relations for means and variances. Let $\bar{X}_n$ and S_n^2 be the mean and variance, respectively, of $X_1, \ldots, X_n$. Then suppose another observation X_{n+1} becomes available. Prove the following:

(a) $\bar{X}_{n+1} = \dfrac{X_{n+1} + n\bar{X}_n}{n+1}$

(b) $nS_{n+1}^2 = (n-1)S_n^2 + \left(\dfrac{n}{n+1}\right)(X_{n+1} - \bar{X}_n)^2$

Answer

(a)

$$\begin{aligned}\bar{X}_{n+1} &= \frac{1}{n+1}\left\{\sum_{i=1}^{n+1} X_i\right\} \\ &= \frac{1}{n+1}\left\{\sum_{i=1}^{n} X_i + X_{n+1}\right\} \\ &= \frac{1}{n+1}\{n\bar{X}_n + X_{n+1}\}\end{aligned}$$

(b) Note: $S_{n+1}^2 = \frac{1}{n}\sum_{i=1}^{n+1}(X_i - \bar{X}_{n+1})^2$ and $S_n^2 = \frac{1}{n-1}\sum_{i=1}^{n}(X_i - \bar{X}_n)^2$

Therefore

$$nS_{n+1}^2 = \sum_{i=1}^{n+1}(X_i - \bar{X}_{n+1})^2$$

$$= \sum_{i=1}^{n+1}\{X_i - \bar{X}_n + \bar{X}_n - \bar{X}_{n+1}\}^2$$

$$= \sum_{i=1}^{n+1}\{(X_i - \bar{X}_n)^2 + 2(X_i - \bar{X}_n)(\bar{X}_n - \bar{X}_{n+1}) + (\bar{X}_n - \bar{X}_{n+1})^2\}$$

$$= \sum_{i=1}^{n+1}(X_i - \bar{X}_n)^2 + 2(\bar{X}_n - \bar{X}_{n+1})\sum_{i=1}^{n+1}(X_i - \bar{X}_n) + (n+1)(\bar{X}_n - \bar{X}_{n+1})^2$$

$$= \sum_{i=1}^{n+1}(X_i - \bar{X}_n)^2 + (X_{n+1} - \bar{X}_n)^2$$

$$+2(\bar{X}_n - \bar{X}_{n+1})\{(n+1)\bar{X}_{n+1} - (n+1)\bar{X}_n\}$$

$$
\begin{aligned}
&\quad +(n+1)(\bar{X}_n - \bar{X}_{n+1})^2 \\
&= (n-1)S_n^2 + (X_{n+1} - \bar{X}_n)^2 - 2(n+1)(\bar{X}_n - \bar{X}_{n+1})^2 \\
&\quad +(n+1)(\bar{X}_n - \bar{X}_{n+1})^2 \\
&= (n-1)S_n^2 + (X_{n+1} - \bar{X}_n)^2 - (n+1)(\bar{X}_n - \bar{X}_{n+1})^2 \\
&= (n-1)S_n^2 + (X_{n+1} - \bar{X}_n)^2 - (n+1)\left\{\bar{X}_n - \tfrac{n\bar{X}_n + X_{n+1}}{n+1}\right\}^2 \\
&= (n-1)S_n^2 + (X_{n+1} - \bar{X}_n)^2 - (n+1)\left\{\tfrac{\bar{X}_n + X_{n+1}}{n+1}\right\}^2 \\
&= (n-1)S_n^2 + (X_{n+1} - \bar{X}_n)^2 - \tfrac{1}{n+1}(\bar{X}_n - X_{n+1})^2 \\
&= (n-1)S_n^2 + \left(1 - \tfrac{1}{n+1}\right)(X_{n+1} - \bar{X}_n)^2 \\
&= (n-1)S_n^2 + \tfrac{n}{n+1}(X_{n+1} - \bar{X}_n)^2.
\end{aligned}
$$