NORMAL, BINORMAL, MULTINORMAL

Jerzy Szulga, *Lectures on Statistics*, Auburn 2002

Notation and writing conventions

The display form of a vector with n components is a (vertical) $n\times 1$ matrix. The transpose of a matrix A is denoted by A^*. Thus:

$$\boldsymbol{x} = \begin{bmatrix} x_1 \\ \cdots \\ x_2 \end{bmatrix} \qquad \boldsymbol{x}^* = [x_1, \ldots, x_n], \quad \boldsymbol{x}^*\boldsymbol{x} = \sum_{k=1}^{n} x_k^2$$

We will use the round-bracket notation in the text, always horizontal. We will use the square-bracket notation in displayed formulas. That is, $\boldsymbol{x} = (x_1, \ldots, x_n)$ appears in the text while $\boldsymbol{x}^*$ in the display. In many texts the transpose is denoted by A^T or A' rather than by A^*.

For a quadratic $n \times n$ matrix $A = [a_{jk}]$, the product $\boldsymbol{x}^* A\, \boldsymbol{x}$ is a quadratic form

$$\boldsymbol{x}^* A\, \boldsymbol{x} = \sum_{j=1}^{n}\sum_{k=1}^{n} a_{jk} x_j x_k.$$

"*Norma*" in Latin means a "carpenter's square"

A standard normal random variable Z has the density

$$\varphi(z) = \frac{1}{\sqrt{2\pi}}\, e^{-\frac{z^2}{2}}, \quad -\infty < z < \infty.$$

The symbol $N(0,1)$ shows that its mean is 0 and the variance is 1.

Let μ and $a \neq 0$ be any numbers. The transformed random variable

$$X = aZ + \mu \qquad \leftrightarrow \qquad Z = \frac{X-\mu}{a} \tag{1}$$

where μ could be interpreted as the translation parameter and $a = \pm\sigma$ (or, $a^2 = \sigma^2$) – the scale parameter, has the density

$$\varphi_X(x) = \frac{1}{\sigma\sqrt{2\pi}} e^{-\frac{(x-\mu)^2}{2\sigma^2}}.$$

We use then the symbol $N(\mu, \sigma^2)$, where $\sigma^2 = a^2$. In other words,

$$X \sim N(\mu, \sigma^2): \qquad \mathsf{E}X = \mu, \qquad \mathsf{Var}\,(X) = \sigma^2 \tag{2}$$

Conversely, if $X \sim N(\mu, \sigma^2)$, then

$$Z = \frac{X - E(X)}{\sqrt{\mathsf{Var}\,(X)}} = \frac{X - \mu}{\sigma} \sim N(0, 1)$$

We often call the above expression **the standardized random variable**.

A **standard bivariate normal** distribution is that of a pair (Z_1, Z_2) of independent $N(0, 1)$ random variables. Therefore, its joint density is the product of the univariate densities:

$$\varphi(z_1, z_2) = \varphi(z_1) \cdot \varphi(z_2) = \frac{1}{2\pi}\, e^{-\frac{z_1^2 + z_2^2}{2}}$$

By the same token, the **standard multivariate normal** density is that of the sequence $(Z_1, Z_2, \ldots, Z_n)$ of independent $N(0, 1)$ random variables, that is

$$\varphi(z_1, z_2, \ldots, z_n) = \varphi(z_1) \cdot \varphi(z_2) \cdots \varphi(z_n) = \frac{1}{(2\pi)^{n/2}}\, e^{-\frac{z_1^2 + z_2^2 + \cdots + z_n^2}{2}} \tag{3}$$

The general bivariate or multivariate normal is an **affine** transformation of the standard bivariate or multivariate normal, analogous to (1). The adjective "affine" means a composition of a translation (adding a number, when $n = 1$) and a linear transformation (scaling, or, multiplication by a scalar, when $n = 1$). That is, if (Z_1, Z_2) is a standard bi-normal, then

$$\begin{aligned} X_1 &= a_{11} Z_1 + a_{12} Z_2 + \mu_1 \\ X_2 &= a_{21} Z_1 + a_{22} Z_2 + \mu_2 \end{aligned} \tag{4}$$

is dubbed the **bivariate normal**.

The relation can be rewritten using the matrix notation. Conventionally, the sequences are written as vertical vectors, so the pairs (X_1, X_2), (Z_1, Z_2) and (μ_1, μ_2) will show as

$$\boldsymbol{X} = \begin{bmatrix} X_1 \\ X_2 \end{bmatrix}, \qquad \boldsymbol{Z} = \begin{bmatrix} Z_1 \\ Z_2 \end{bmatrix}, \qquad \boldsymbol{\mu} = \begin{bmatrix} \mu_1 \\ \mu_2 \end{bmatrix}$$

The scaling coefficients can be noted as the 2×2 matrix

$$A = \begin{bmatrix} a_{11} & a_{12} \\ a_{21} & a_{22} \end{bmatrix},$$

and the affine transformation (4) becomes a matrix formula analogous to (1):

$$\boldsymbol{X} = A\boldsymbol{Z} + \boldsymbol{\mu} \tag{5}$$

Note that the same formula connotes the n**-variate normal** $(X_1, ...X_n)$:

$$\begin{aligned} X_1 &= a_{11}Z_1 + a_{12}Z_2 + \cdots + a_{1n}Z_n + \mu_1 \\ &\cdots\cdots\cdots \\ X_n &= a_{n1}Z_1 + a_{n2}Z_2 + + \cdots + a_{nn}Z_n + \mu_n, \end{aligned} \tag{6}$$

where now

$$\boldsymbol{X} = \begin{bmatrix} X_1 \\ \cdots \\ X_n \end{bmatrix}, \qquad A = \begin{bmatrix} a_{11} & \cdots & a_{1n} \\ \cdots & \cdots & \cdots \\ a_{n1} & \cdots & a_{nn} \end{bmatrix}, \qquad \boldsymbol{Z} = \begin{bmatrix} Z_1 \\ \cdots \\ Z_n \end{bmatrix}, \qquad \boldsymbol{\mu} = \begin{bmatrix} \mu_1 \\ \cdots \\ \mu_n \end{bmatrix}.$$

In the matrix notation, the standard normal density (3) can be written as follows. We write $\exp\{\cdots\} = e^{\cdots}$, and A^* denotes the transpose of the matrix A. Thus, for example, if $\boldsymbol{z}$ is a vertical $n \times 1$ vector, then $\boldsymbol{z}^*$ is the horizontal $1 \times n$ vector. In particular

$$\boldsymbol{z}^*\boldsymbol{z} = z_1^2 + z_2^2 + \cdots + z_n^2.$$

Hence, (3) becomes

$$\varphi(\boldsymbol{z}) = \frac{1}{(2\pi)^{n/2}}\, exp\left\{ -\frac{\boldsymbol{z}^*\boldsymbol{z}}{2} \right\}$$

To find the joint density of the transformed normal (5) (which corresponds to (4) or (6)), we could use the change-of-variable formula (page 98) involving the Jacobian of the transformation (4) (or (6))

$$J = \left| \frac{\partial \boldsymbol{x}}{\partial \boldsymbol{z}} \right| = |\det A|\,,$$

But, it is allowed only when the transformation is one-to-one, that is, the matrix A has the inverse

$$A^{-1} = \begin{bmatrix} c_{11} & \cdots & c_{1n} \\ \cdots & \cdots & \cdots \\ c_{n1} & \cdots & c_{nn}, \end{bmatrix}$$

In this case, (5) can be rewritten as

$$\boldsymbol{Z} = A^{-1}\Big(\boldsymbol{X} - \boldsymbol{\mu}\Big)$$

So, assume that A has the inverse, which happens if and only if $\det A \neq 0$. Then

$$\varphi_{\boldsymbol{X}}(\boldsymbol{x}) = \frac{\varphi\Big(A^{-1}(\boldsymbol{x}-\boldsymbol{\mu})\Big)}{|\det A|} = \frac{\exp\left\{-\frac{(\boldsymbol{x}-\boldsymbol{\mu})^* A^{-1*} A^{-1}(\boldsymbol{x}-\boldsymbol{\mu})}{2}\right\}}{|\det A| \cdot (2\pi)^{n/2}} \tag{7}$$

Example. 1. Consider the bivariate normal

$$\begin{aligned} X_1 &= Z_1 - 4Z_2 + 7 \\ X_2 &= 2Z_1 + 3Z_2 - 5 \end{aligned} \qquad \Rightarrow A = \begin{bmatrix} 1 & -4 \\ 2 & 3 \end{bmatrix} \qquad \boldsymbol{\mu} = \begin{bmatrix} 7 \\ -5 \end{bmatrix}.$$

For 2×2 matrices, the inverse is easy to find:

1) swap diagonal elements
2) change the sign at the off-diagonal elements
3) divide by the determinant

Warning:
<u>Never use this method for $n > 2$</u>! Use the Gauss-Jordan row reduction method instead.

Hence,

$$\det(A) = 11, \qquad A^{-1} = \frac{1}{11}\begin{bmatrix} 3 & 4 \\ -2 & 1 \end{bmatrix}.$$

Then

$$\boldsymbol{Z} = A^{-1}(\boldsymbol{x}-\boldsymbol{\mu}) = \frac{1}{11}\begin{bmatrix} 3 & 4 \\ -2 & 1 \end{bmatrix} \cdot \begin{bmatrix} x_1 - 7 \\ x_2 + 5 \end{bmatrix} = \frac{1}{11}\begin{bmatrix} 3x_1 + 4x_2 - 1 \\ -2x_1 + x_2 + 19 \end{bmatrix}$$

Therefore, (7) reads

$$f_{X_1,X_2}(x_1,x_2) = \frac{1}{11} \cdot \frac{1}{2\pi} \exp\left\{-\frac{(3x_1+4x_2-1)^2 + (-2x_1+x_2+19)^2}{11^2 \cdot 2}\right\}.$$

This method fails when the transformation is not one-to-one (one would have 0 in the denominator). In fact, the density will not exist at all, regardless of any method. For example, the random pair $(Z_1 + Z_2, 2Z_1 + 2Z_2)$ does not have a density although it is not discrete.

The covariance approach

Given transformation (4), we can compute variances and covariances of the transformed variables:

$$\mathsf{E}(X_1) = \mu_1; \quad \mathsf{E}(X_2) = \mu_2;$$
$$\mathsf{Var}\,(X_1) = a_{11}^2 + a_{21}^2; \quad \mathsf{Var}\,(X_2) = a_{12}^2 + a_{22}^2; \quad \mathsf{Cov}\,(X_1, X_2) = a_{11}a_{21} + a_{12}a_{22}.$$

Note that

$$\begin{bmatrix} a_{11} & a_{12} \\ a_{21} & a_{22} \end{bmatrix} \begin{bmatrix} a_{11} & a_{21} \\ a_{12} & a_{22} \end{bmatrix} = A\,A^*, \tag{8}$$

where A^* denotes the transpose of A (sometimes also denoted by A^T). In other words, the elements of the matrix AA^* are covariances, so $\Sigma = AA^*$ is called the **covariance matrix** or, simply, **variance** (because it plays the role of the variance in formula (2)). Similarly we handle the n-variate case. Summarizing the above in the matrix notation we obtain almost a copy of (2)

$$\boldsymbol{X} \sim N(\boldsymbol{\mu}, \Sigma): \qquad \mathsf{E}\boldsymbol{X} = \boldsymbol{\mu}; \qquad \mathsf{Var}\,(\boldsymbol{X}) = A\,A^* \stackrel{\mathsf{df}}{=} \Sigma. \tag{9}$$

The notation $N(\boldsymbol{\mu}, \Sigma)$ is correct if we know that the probability distribution of the random sequence $\boldsymbol{X}$ depends solely on means and covariances. But we don't know that, yet. We only know that the distribution depends on coefficients a_{ij}.

So, we have to discuss the issue: let $\boldsymbol{X} = (X_1, \ldots, X_n)$ be a multivariate normal, i.e., a sequence of two or more variables. Let $\boldsymbol{\mu}$ be a sequence of means, and let Σ be the matrix of covariances. That is

$$\Sigma = \begin{bmatrix} s_{11} & \cdots & s_{1n} \\ \cdots & \cdots & \cdots \\ s_{n1} & \cdots & s_{nn} \end{bmatrix}, \qquad \text{where} \;\; s_{ij} = \mathsf{Cov}\,(X_i, X_j), \quad s_{ii} = \mathsf{Cov}\,(X_i, X_i) = \mathsf{Var}\,(X_i).$$

If we show that and how the presumed coefficients a_{ij} can be found, i.e., if we can solve for A the matrix equation

$$\Sigma = A\,A^*, \tag{10}$$

then we solve the problem. Notice that in the univariate case we're solving

$$\sigma^2 = a^2,$$

which is easy but has two solutions: $a = \sigma$ and $a = -\sigma$. This means only that the random variables

$$aZ + \mu \qquad \text{and} \qquad -aZ + \mu$$

have the same probability distribution, which is not surprising because the standard normal Z is symmetric (its density is an even function, $\varphi(z) = \varphi(-z)$).

When $n = 2$ we can rewrite (8) explicitly and easily. Since the variance matrix is symmetric, hence we obtain three equations with four unknowns a_{ij}:

$$\begin{bmatrix} a_{11}^2 + a_{21}^2 & a_{11}a_{21} + a_{12}a_{22} \\ a_{11}a_{21} + a_{12}a_{22} & a_{12}^2 + a_{22}^2 \end{bmatrix} = \begin{bmatrix} s_{11} & s_{12} \\ s_{12} & s_{22} \end{bmatrix}$$

In the n-variate case, we search for n^2 uknowns a_{ij} but – because of the symmetry of Σ – the equations related to the lower triangle (under the diagonal) simply repeat the equations related to the upper triangle (above the diagonal). Hence, the number of equations is (the diagonal's size + the upper /or lower/ triangle's size):

$$n + \frac{n(n-1)}{2} = \frac{n(n+1)}{2}$$

If A solves (10) and Q is any $n \times n$ matrix with the property

$$QQ^* = I \quad (I \text{ is the identity matrix, so } Q^* = Q^{-1}), \tag{11}$$

then AQ also solve (10). Matrices satisfying (11) are called **orthogonal**. Their rows (or columns) are mutually orthogonal unit vectors in $\mathbb{R}^n$. Accidentally, we have proved that

An orthogonal transformation $\boldsymbol{Y} = Q\boldsymbol{Z}$ preserves the distribution.

For example,

$$Q = \begin{bmatrix} \frac{1}{\sqrt{2}} & \frac{1}{\sqrt{2}} \\ \frac{1}{\sqrt{2}} & \frac{-1}{\sqrt{2}} \end{bmatrix}, \quad Y_1 = \frac{Z_1 + Z_2}{\sqrt{2}}, \; Y_2 = \frac{Z_1 - Z_2}{\sqrt{2}}.$$

In particular, when Z_1, Z_2 are independent standard normal, so are Y_1, Y_2. Although it may seem surprising that, e.g., $Z_1 + Z_2$ and $Z_1 - Z_2$ are independent, the application of an orthogonal matrix to the joint density amounts to the revolution or symmetry of its graph, a bell shaped surface, about the z-axis. But since the surface is already a symmetric surface of revolution, no more revolution or symmetry will have any effect.

So, how to solve equation (10)?

Method 1. By hand. Not recommended.

Method 2. Using a calculator or software.

More expensive calculators or computational software like Matlab, Maple, Mathematica, etc., may have the needed command to perform – so called – **Cholesky factorization**

Example. Consider the covariance matrix

$$S = \Sigma = \begin{bmatrix} 1 & 2 & 3 \\ 2 & 5 & 7 \\ 3 & 7 & 26 \end{bmatrix}$$

Matlab `S=[1,2,3;2,5,7;3,7,26]; A=chol(S)`*

Maple `with(linalg): S := matrix(3,3,[1,2,3,2,5,7,3,7,26]); A := cholesky(S);`

return the lower triangular matrix:

$$\mathrm{A} = \begin{bmatrix} 1 & 0 & 0 \\ 2 & 1 & 0 \\ 3 & 1 & 4 \end{bmatrix}$$

The obtained reult means that the 3-variate normal vector $(Z_1,\ 2Z_2+Z_3,\ 3Z_1+Z_2+4Z_3)$ has the covariance matrix S.

Since $\det(A) = \det(A^*)$, so $\det(A) = \det(\Sigma)^{1/2}$, then the general density (7) now can be written

$$\begin{aligned} \varphi_{\boldsymbol{X}}(\boldsymbol{x}) &= \frac{1}{\det(\Sigma)^{1/2}\cdot(2\pi)^{n/2}} \exp\left\{ \frac{(\boldsymbol{x}-\boldsymbol{\mu})^*\,\Sigma^{-1}\,(\boldsymbol{x}-\boldsymbol{\mu})}{2} \right\} \\ &= \frac{1}{\det(\Sigma)^{1/2}\cdot(2\pi)^{n/2}} \exp\left\{ -\frac{1}{2}\sum_{i=1}^{n}\sum_{j=1}^{n} c_{ij}(x_i-\mu_i)(x_j-\mu_j) \right\}. \end{aligned}$$

Since the moment generating function of the standard density is

$$M(\boldsymbol{t}) = \mathsf{E}\exp\{\,\boldsymbol{t}^*\boldsymbol{Z}\,\} = \exp\left\{\frac{\boldsymbol{t}^*\boldsymbol{t}}{2}\right\}$$

and

$$M_{\boldsymbol{X}}(\boldsymbol{t}) = \mathsf{E}\exp\{\,\boldsymbol{t}^*\boldsymbol{X}\,\} = \mathsf{E}\exp\{\,\boldsymbol{t}^*(A\boldsymbol{Z}+\boldsymbol{\mu}\,\} = \mathsf{E}\exp\{\,\boldsymbol{t}^*A\boldsymbol{Z}\,\}\,\exp\{\,\boldsymbol{t}^*\boldsymbol{\mu}\,\}$$
$$= \exp\{\,\boldsymbol{t}^*AA^*\boldsymbol{t}/2\,\}\,\exp\{\,\boldsymbol{t}^*\boldsymbol{\mu}\,\}\,,$$

hence the m.g.f. in general is

$$\begin{aligned} M_{\boldsymbol{X}}(\boldsymbol{t}) &= \exp\left\{\frac{1}{2}\boldsymbol{t}^*\Sigma\boldsymbol{t}+\boldsymbol{t}^*\boldsymbol{\mu}\right\} \\ &= \exp\left\{\frac{1}{2}\sum_{i=1}^n\sum_{j=1}^n\sigma_{ij}\,t_i\,t_j+\sum_{i=1}^n t_jx_j\right\}. \end{aligned}$$

In our example, Matlab returns

```
inv(S)


ans =


    5.0625   -1.9375   -0.0625
   -1.9375    1.0625   -0.0625
   -0.0625   -0.0625    0.0625
```

That is, the quadratic form is (doubling the off-diagonal elements, since it is a symmetric matrix, so e.g., combining $-1.9375x_1x_2 - 1.9375x_2x_1$ into $-2 * 1.9375\,x_1x_2$)

$$(\boldsymbol{x}-\boldsymbol{\mu})^*\,\Sigma^{-1}\,(\boldsymbol{x}-\boldsymbol{\mu}) = 5.0625x_1^2-3.875x_1x_2-0.125x_1x_3+2.125x_2^2-0.125x_2x_3+0.0625x_3^2$$

It is especially easy to find the inverse of a 2×2 matrix:

$$\begin{aligned} \Sigma^{-1} = \begin{bmatrix} \sigma_1^2 & \sigma_{12} \\ \sigma_{12} & \sigma_2^2 \end{bmatrix}^{-1} &= \frac{1}{\sigma_1^2\sigma_2^2-\sigma_{12}^2} \begin{bmatrix} \sigma_2^2 & -\sigma_{12} \\ -\sigma_{12} & \sigma_1^2 \end{bmatrix} \\ &= \frac{\sigma_1^2\sigma_2^2}{\sigma_1^2\sigma_2^2-\sigma_{12}^2} \begin{bmatrix} \frac{\sigma_2^2}{\sigma_1^2\sigma_2^2} & \frac{-\sigma_{12}}{\sigma_1^2\sigma_2^2} \\ \frac{-\sigma_{12}}{\sigma_1^2\sigma_2^2} & \frac{\sigma_1^2}{\sigma_1^2\sigma_2^2} \end{bmatrix} \\ &= \frac{1}{1-\rho^2} \begin{bmatrix} \frac{1}{\sigma_1^2} & \frac{-\rho}{\sigma_1\sigma_2} \\ \frac{-\rho}{\sigma_1\sigma_2} & \frac{1}{\sigma_2^2} \end{bmatrix} \end{aligned}$$

SUMMARY

A random vector $\boldsymbol{Z} = (Z_1, \ldots, Z_n)$ is called **standard normal**, $N(\mathbf{0}, I)$ in short, if its components are i.i.d. $N(0,1)$ variables. In other word, its density φ and the m.g.f. $M(t)$ are as follows:

$$\varphi(\boldsymbol{z}) = \frac{1}{(2\pi)^{n/2}}\, exp\left\{-\frac{\boldsymbol{z}^*\boldsymbol{z}}{2}\right\}, \qquad M(t) = \mathsf{E}\exp\{\,\boldsymbol{t}^* Z\,\} = \exp\left\{\frac{\boldsymbol{t}^*\boldsymbol{t}}{2}\right\}.$$

A general n**-variate normal distribution** $N(\boldsymbol{\mu}, \Sigma)$, or a n**-dimensional normal vector** $\boldsymbol{X} = (X_1, \ldots, X_n)$ where

$$\mu_k = \mathsf{E}X_k, \qquad \Sigma = [s_{jk}], \quad s_{jk} = \mathsf{Cov}\,(X_j, X_k),$$

may be introduced in one of the following three ways

the density $\qquad \varphi_{\boldsymbol{X}}(\boldsymbol{x}) = \dfrac{1}{\det(\Sigma)^{1/2}\cdot(2\pi)^{n/2}}\exp\left\{\dfrac{(\boldsymbol{x}-\boldsymbol{\mu})^*\,\Sigma^{-1}\,(\boldsymbol{x}-\boldsymbol{\mu})}{2}\right\},$

the m.g.f. $\qquad M_{\boldsymbol{X}}(\boldsymbol{t}) = \mathsf{E}\exp\{\,\boldsymbol{t}^*\boldsymbol{X}\,\} = \mathsf{E}\exp\left\{\dfrac{1}{2}\boldsymbol{t}^*\Sigma\,\boldsymbol{t} + \boldsymbol{t}^*\boldsymbol{\mu}\right\},$

a transformed $N(\mathbf{0}, I)$ $\qquad \boldsymbol{X} = A\boldsymbol{Z} + \boldsymbol{\mu}$, where $\Sigma = A\,A^*$.

Immediate properties. An orthogonal matrix Q is such that its rows (or columns) are orthogonal unit vectors. In particular, its transpose equals its inverse, or, $QQ^* = I$.

1. *An orthogonal transformation of a standard normal yields a standard normal vector.*

   The transformation method #3 (i.e., the matrix A) is not unique because its replacement by the product AQ, where Q is an orthogonal matrix, yields $(AQ)(AQ)^* = AQQ^*A^* = AIA^* = AA^*$, i.e., has no effect on the m.g.f.

2. *Two normal random variables are independent if and only if they are uncorrelated.*

   The independence always implies the zero covariance. Conversely, if the covariance is 0, then the 2×2 covariance matrix of (X_1, X_2) is diagonal. Then, the inverse (used in the density) is diagonal as well. So, the joint density can be written as the product of the univariate densities. Similarly, the m.g.f. can be written as the product of the univariate m.g.f.'s.

3. *Let $\boldsymbol{Z} = (Z_1, \ldots, Z_n)$ be n-variate standard normal. Let s_k be any sequence of scalars. Then*

$$\overline{Z} = \frac{Z_1 + \cdots + Z_n}{n} \quad \textit{and} \quad \sum_{k=1}^{n} s_k \Big(Z_k - \overline{Z} \Big) \quad \textit{are independent.}$$

It is enough to see that the covariance is 0. Using the algbera of expectations and variances,

$$\mathsf{E}\left(\overline{Z} \cdot \sum_{k=1}^{n} s_k \Big(Z_k - \overline{Z} \Big) \right) = \sum_{k=1}^{n} s_k \Big(\mathsf{E}(\overline{Z}\, Z_k) - \mathsf{E}(\overline{Z})^2 \Big) = \sum_{k=1}^{n} s_k \Big(\frac{1}{n} - \frac{1}{n} \Big) = 0$$

4. *Let $\boldsymbol{Z}$ be n-variate standard normal. Then the sample mean $\overline{Z}$ and the sample variance $\frac{1}{n} \sum_{k=1}^{n} \Big(Z_k - \overline{Z} \Big)^2$ are independent.*

If x_k are scalars and S_k are ± 1 Bernoulli, $\mathsf{P}(S_k = 1) = \mathsf{P}(S_k = -1) = 1/2$, then

$$\sum_{k=1}^{n} x_k^2 = \mathsf{Var}\Big(\sum_{k=1}^{n} S_k x_k \Big) = \mathsf{E}\Big| \sum_{k=1}^{n} S_k x_k \Big|^2 = \frac{1}{2^n} \Big| \sum_{\boldsymbol{s}} s_k x_k \Big|^2$$

where the last sum of size 2^n is taken over all 2^n choices of signs $s_k = \pm 1$.

This shows that the sample variance can be obtained by algebraic operations (squaring, adding, multiplying by constants) peformed on random variables $\sum_k s_k (Z_k - \overline{Z})$ which remain independent of $\overline{X}$. Hence, the outcome which is the sample variance remains independent.