Square roots. Recall the Spectral Theorem of linear algebra: If $M \in \mathbb{R}^{n\times n}$ (i.e., M is an $n \times n$ matrix whose entries are real) and $M' = M$ (i.e., M is symmetric) then there exists a matrix $G \in \mathbb{O}_n$ (i.e., G is an $n \times n$ orthogonal matrix; "G is orthogonal" means the entries in G are real and $GG' = G'G = I$) and there exist real numbers $\lambda_1, \ldots, \lambda_n$ such that

$$M = G \begin{bmatrix} \lambda_1 & & \\ & \ddots & \\ & & \lambda_n \end{bmatrix} G'.$$

If $\lambda_1, \ldots, \lambda_n \geq 0$ then we can write

$$M^{1/2} = G \begin{bmatrix} \sqrt{\lambda_1} & & \\ & \ddots & \\ & & \sqrt{\lambda_n} \end{bmatrix} G'$$

and then we have

$$\begin{aligned} M^{1/2}M^{1/2} &= G \begin{bmatrix} \sqrt{\lambda_1} & & \\ & \ddots & \\ & & \sqrt{\lambda_n} \end{bmatrix} G'G \begin{bmatrix} \sqrt{\lambda_1} & & \\ & \ddots & \\ & & \sqrt{\lambda_n} \end{bmatrix} G' \\ &= G \begin{bmatrix} \sqrt{\lambda_1} & & \\ & \ddots & \\ & & \sqrt{\lambda_n} \end{bmatrix} \begin{bmatrix} \sqrt{\lambda_1} & & \\ & \ddots & \\ & & \sqrt{\lambda_n} \end{bmatrix} G' \\ &= G \begin{bmatrix} \lambda_1 & & \\ & \ddots & \\ & & \lambda_n \end{bmatrix} G' = M, \end{aligned}$$

and so our notation is justified: What we have called "$M^{1/2}$" really is a square root of M.

Square roots of variances. Now suppose X is an $n \times 1$ random vector, $\mu = \mathbf{E}(X)$, and $M = \mathbf{var}(X) = \mathbf{E}((X-\mu)(X-\mu)')$. If we define $W = [W_1, \ldots, W_n]' = G'X$ then

$$\mathbf{var}(W) = G'(\mathbf{var}(X))G = G'MG = G' \left(G \begin{bmatrix} \lambda_1 & & \\ & \ddots & \\ & & \lambda_n \end{bmatrix} G' \right) G = \begin{bmatrix} \lambda_1 & & \\ & \ddots & \\ & & \lambda_n \end{bmatrix}.$$

This implies $\lambda_i = \mathbf{var}(W_i)$ and therefore λ_i must be ≥ 0. Consequently $M = \mathbf{var}(X)$ has a square root $M^{1/2}$ of the sort we saw above.

We can see that that square root has an inverse if $\lambda_i \neq 0$ for each i, by showing that

$$M^{-1/2} = G \begin{bmatrix} \lambda_1^{-1/2} & & \\ & \ddots & \\ & & \lambda_n^{-1/2} \end{bmatrix} G'$$

is that inverse. Just multiply this by the diagonalized form of $M^{1/2}$ that we saw on the previous page, observe the cancellation of "$G'G$" in the middle, etc.

<u>**Central Limit Theorem.**</u> Now imagine $X_1, X_2, X_3, \ldots \sim$ i. i. d. with expectation $\mathbf{E}(X_1) = \mu \in \mathbb{R}^n$ and variance $\mathbf{var}(X_1) = M \in \mathbb{R}^{n\times n}$, but <u>do not</u> assume X_1 is normally distributed. Then, as $n \to \infty$, the probability distribution of

$$(M/n)^{-1/2}\left(\frac{X_1 + \cdots + X_n}{n} - \mu\right)$$

approaches $N_n(0, I_n)$.

Instead of dividing by the standard deviation, we have multiplied by $(M^{-1/2}/n)$, where M/n is the variance of $(X_1 + \cdots + X_n)/n$. If $n = 1$, that is the same as dividing by the standard deviation, but this version still works when $n > 1$.

If $Z \sim N_n(0, I_n)$ then $\|Z\|^2 \sim \chi^2_n$.

But what if some of the λs are zero? In that case let us say

$$M^{-1/2} = G \begin{bmatrix} \lambda_1^{-1/2} & & & & & \\ & \ddots & & & & \\ & & \lambda_k^{-1/2} & & & \\ & & & 0 & & \\ & & & & \ddots & \\ & & & & & 0 \end{bmatrix} G'.$$

This is not really an inverse, but rather a sort of generalized inverse. Then the limiting distribution in our Central Limit Theorem, rather than $N_n(0, I_n)$, will be

$$N_n\left(0, G \begin{bmatrix} 1 & & & & & \\ & \ddots & & & & \\ & & 1 & & & \\ & & & 0 & & \\ & & & & \ddots & \\ & & & & & 0 \end{bmatrix} G'\right)$$

where there are k 0s and $n-k$ 1s.

If Z is normally distributed with expecation 0 and variance equal to the above matrix, then

$$G'Z \sim N_n\left(0, \begin{bmatrix} 1 & & & & & \\ & \ddots & & & & \\ & & 1 & & & \\ & & & 0 & & \\ & & & & \ddots & \\ & & & & & 0 \end{bmatrix}\right)$$

and therefore $\|G'Z\|^2 \sim \chi^2_k$.

But how is $\|Z\|^2$ distributed?

$$\|G'Z\|^2 = (G'Z)'(G'Z) = (Z'G)(G'Z) = Z'(GG')Z = Z'Z = \|Z\|^2.$$

So we also have $\|Z\|^2 \sim \chi^2_k$.

Back to the problem of throwing dice. A "fair" die is thrown n times. Let X_i = the number of times i appears, for $i = 1, 2, 3, 4, 5, 6$. We found that $\mathbf{var}(X) = (n/6)Q_6$ where Q_6 is the 6×6 matrix of rank 5 characterized by

$$Q_6 \begin{bmatrix} u_1 \\ u_2 \\ u_3 \\ u_4 \\ u_5 \\ u_6 \end{bmatrix} = \begin{bmatrix} u_1 - \overline{u} \\ u_2 - \overline{u} \\ u_3 - \overline{u} \\ u_4 - \overline{u} \\ u_5 - \overline{u} \\ u_6 - \overline{u} \end{bmatrix} \quad \text{where } \overline{u} = \frac{u_1 + u_2 + u_3 + u_4 + u_5 + u_6}{6}.$$

Our Central Limit result says that if n is big then the distribution of

$$(Q_6/(6n))^{-1/2}\left(\frac{X}{n} - \frac{1}{6}\begin{bmatrix} 1 \\ 1 \\ 1 \\ 1 \\ 1 \\ 1 \end{bmatrix}\right)$$

is approximately normal with expectation 0 and variance

$$G\begin{bmatrix} 1 & & & & & \\ & 1 & & & & \\ & & 1 & & & \\ & & & 1 & & \\ & & & & 1 & \\ & & & & & 0 \end{bmatrix}G'$$

for some orthogonal matrix G. It follows that

$$6n\left\| Q_6^{-1}\left(\frac{X}{n}-\frac{1}{6}\begin{bmatrix}1\\1\\1\\1\\1\\1\end{bmatrix}\right)\right\|^2 \sim \chi_5^2$$

approximately, if n is big enough.

To simplify this, first observe that Q_6 is its own generalized inverse, because we're just inverting those 1s on the diagonal. Then observe that

$$Q_6\left(\frac{X}{n}-\frac{1}{6}\begin{bmatrix}1\\1\\1\\1\\1\\1\end{bmatrix}\right)=\left(\frac{X}{n}-\frac{1}{6}\begin{bmatrix}1\\1\\1\\1\\1\\1\end{bmatrix}\right)$$

because the components of X must add up to n, so the components of X/n must add up to 1, so the components of the difference must add up to 0, hence, subtracting their average from each of them is just subtracting 0. The we get

$$6n\left\|\frac{X}{n}-\frac{1}{6}\begin{bmatrix}1\\1\\1\\1\\1\\1\end{bmatrix}\right\|^2=\sum_{i=1}^{6}\frac{(X_i-n/6)^2}{n/6}$$

so this is distributed as χ_5^2, provided the null hypothesis of "fairness" of the die is assumed.