Additional remarks on the four assumptions.

Note that under Assumption 4, we have that

$$u_i \sim N(0, \sigma^2), \text{ for all } i.$$

This assumption clearly implies that

$$E(u_i) = 0, \text{ for all } i \text{ (Assumption 1)},$$

and that

$$Var(u_i) = \sigma^2, \text{ for all } i \text{ (Assumption 2)}.$$

That is, the notation "$N(0,\sigma^2)$" <u>means</u>, among other things, that the random variable in question has mean (or expected value) of zero, and a variance of σ^2.

If we augment Assumption 4 to state that the $\{u_i\}$ errors are drawn <u>independently</u> from a $N(0,\sigma^2)$ distribution — we already are assuming that they are drawn from an <u>identical</u> distribution — then the modified Assumption 4 also implies Assumption 3, namely that

$$E(u_i u_j) = 0 \text{ for } i \neq j.$$

The right way to state Assumption 4 is thus to say that all the errors are identically and independently drawn ('i.i.d.') from a Normal $(0, \sigma^2)$ distribution. Or, more compactly,

$$u_i \sim \text{i.i.d. } N(0, \sigma^2), \text{ for all } i.$$

The logic is that if error terms are independently drawn, then they are necessarily uncorrelated with one another — any observed correlation would be purely by chance.

Note well: if we state the modified Assumption 4, then this combines (in the sense that it implies) all three of our earlier assumptions, and this is economical.

Against this, Assumption 4 commits us to a <u>particular</u> distribution, and that is a very strong assumption — much stronger than Assumptions 1–3.

The chief gain is that, if we want to conduct hypothesis testing, then we need this kind of distributional assumption.