

general aspect of the crop was still decidedly green, and the last one when about half a bushel of grain per acre had been shaken on to the ground, so that the various cuttings embraced the whole period during which any one would be likely to harvest a crop on a farming scale. Complications possibly arising from differences in weather after cutting were reckoned with in a subsidiary experiment. The actual stage of maturity at which each cutting was made will be described later.

Before proceeding to state the weights of the 100 grains cut at the various stages, it is perhaps necessary to explain why so large a number as thirty plots were taken from which to cut representative straws at each stage. It was because of the generally unsatisfactory nature of averages as commonly determined in agricultural experiments. An average alone gives a scanty amount of information. "An average of 20" may mean that there were 100 observations between 19 and 21, with a mean of 20; or it may mean that there were two observations of 1 and 39. Again, in tossing 20 coins 20 times we may find that an average of 9 heads will turn up—a result that is obviously not strictly reliable. Therefore mathematicians have invented a device to indicate the reliability of any average. They add to the figure indicating the mean another figure indicating what is called the "probable error." Thus, if an average is stated to be  $20 \pm 1$  the 1 is called the probable error, meaning that if another average were computed with equal care the chances would be even that the new average would be between 19 and 21, or outside these limits. Not the clearest of indications, one might think, but one that becomes quite easily understandable with use. It is clear that the smaller the probable error the more reliable the average.

Returning to the Lincoln experiment, the twelve heads of each of the thirty A plots and the twelve from each of the thirty B plots, &c., were hung up for a couple of months until they were all dry, and assumed to have the same water content—an assumption that was checked and found correct later on. They were all threshed on the one day, and from each of the 154 lots 100 grains were counted and weighed within two days. The averages with their probable errors were as follows:—

*Average Weight (from thirty or thirty-one Plots) of 100 Grains of Solid-straw Tuscan Wheat cut at Three-day Intervals.*

Stage A (the earliest)	..	4.047 $\pm$ 0.024	grammes.
" B cut three days later	..	4.060 $\pm$ 0.026	"
" C	..	4.299 $\pm$ 0.028	"
" D	..	4.269 $\pm$ 0.032	"
" E	..	4.299 $\pm$ 0.045	"

These figures are of considerable interest. The smallness of the probable errors (0.5 to 1.0 per cent. of the averages) is satisfactory proof that the weighings were accurate and their number sufficient, but no explanation occurs to us of their gradual increase with the advance in the date of cutting. There is a general similarity between the A and B cuttings, then a distinct break between B and C, and then a similarity between C, D, and E. The variations between the last three—*i.e.*, the falling-away at D to less than C or E—might