



# THEOREM OF THE DAY

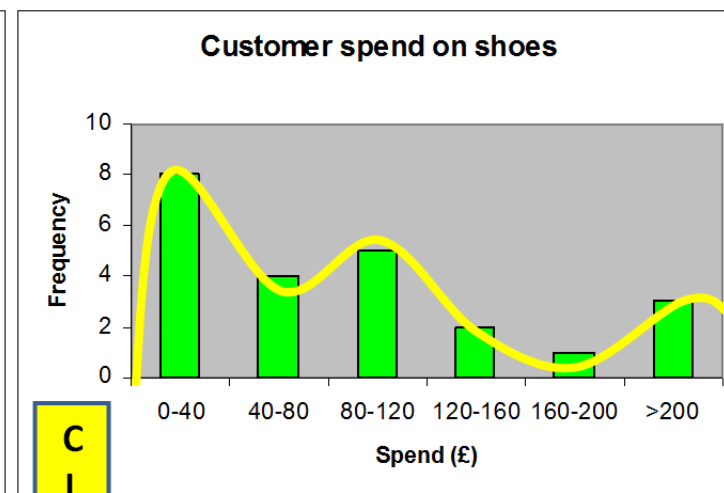
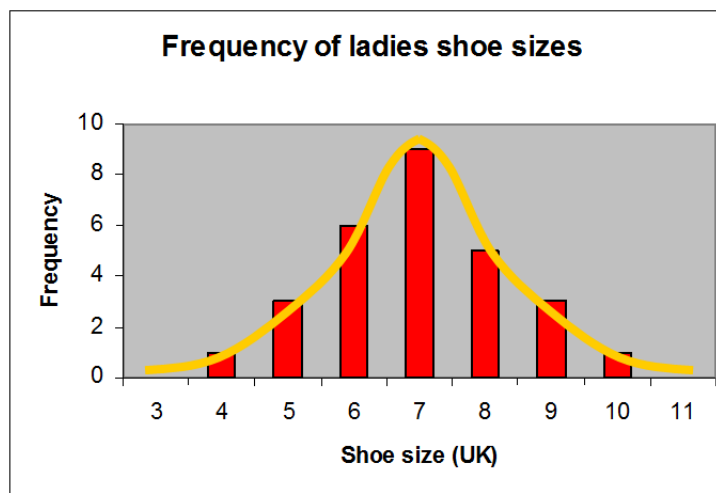
**The Central Limit Theorem** Suppose  $S_1, S_2, S_3, \dots$  are sets of size  $n$  obtained by sampling, with replacement, from some distribution having finite mean  $\mu$  and finite nonzero variance  $\sigma^2$ . Then the sample means  $s_1, s_2, s_3, \dots$  are asymptotically (as  $n \rightarrow \infty$ ) normally distributed with mean  $\mu$  and variance  $\sigma^2/n$ .

Data which is normally distributed (e.g. shoe size, following a symmetrical, bell-shaped curve) allows accurate probability estimation (how many size 4 shoes will you sell, etc). Data that is unhelpfully distributed (how much each customer spends) is much harder to analyse. The Central Limit Theorem says that if we average over each day's customers, a normal distribution will magically reappear. Better still, if we take weekly averages: the larger the sample, the stronger the effect.

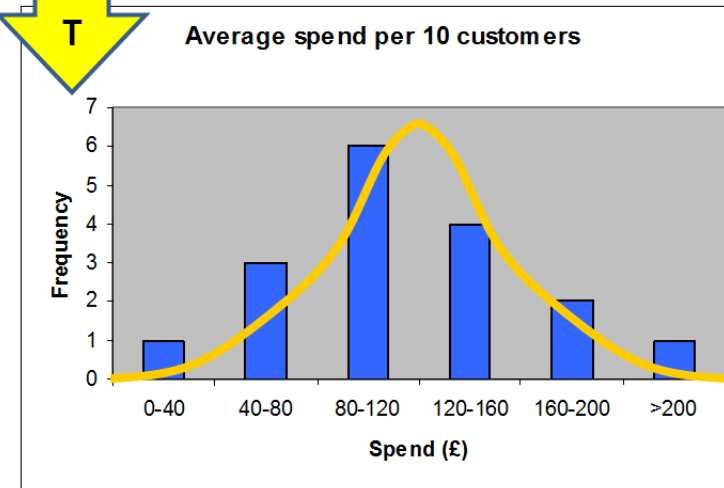
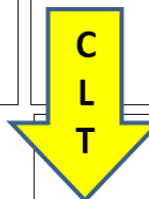
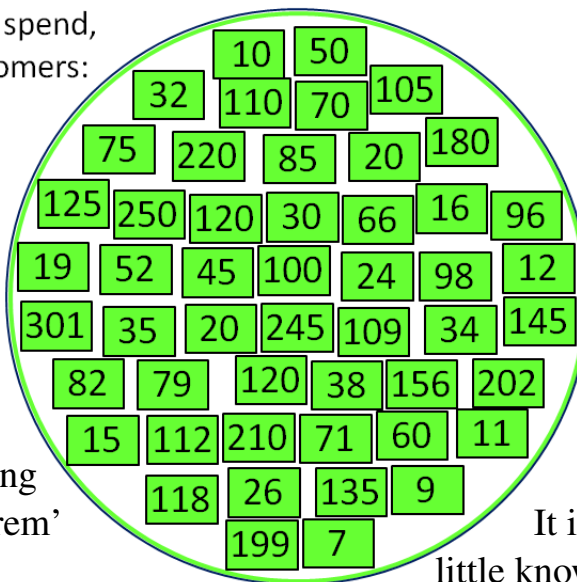
Try this for yourself: the circle contains 50 values randomly distributed in the proportions indicated by the customer spend distribution on the far right. Closing your eyes, touch the circle with a pen and choose the nearest number. Do this 10 times for a random sample and plot the mean, to the nearest whole number, on a frequency chart. If you repeat this often enough, the bar heights will eventually 'settle down' to correspond to a normal distribution, e.g. bottom left.

The central limit theorem can be traced back to De Moivre in 1738, but the rigorous development of its theory was due to Russian mathematicians, particularly Chebyshev, Markov and Lyapunov, culminating in formal proofs at the beginning of the 20th century. The name 'Central Limit Theorem' (in German) is due to George Pólya in 1920.

It is interesting to note that probability theory was so little known in England at this time that Alan Turing's 'discovery' and proof of the Central Limit Theorem in 1935 was enough to get him elected to a fellowship of King's College Cambridge. His proof was essentially that given by Lindeberg in 1922 (Zabell, S., *Amer. Math. Monthly*, 102, 483–494).



Sample spend, 50 customers:



**Web link:** [www2.sjsu.edu/faculty/watkins/randovar.htm](http://www2.sjsu.edu/faculty/watkins/randovar.htm)

**Further reading:** *The Life and Times of the Central Limit Theorem* by William J. Adams, American Mathematical Society, 2010.

