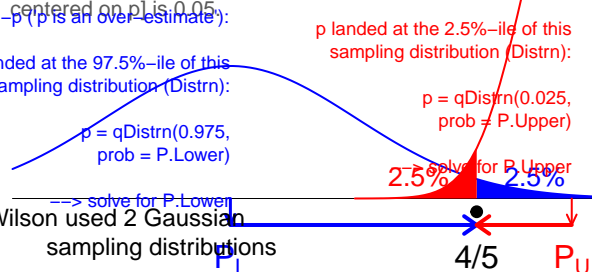


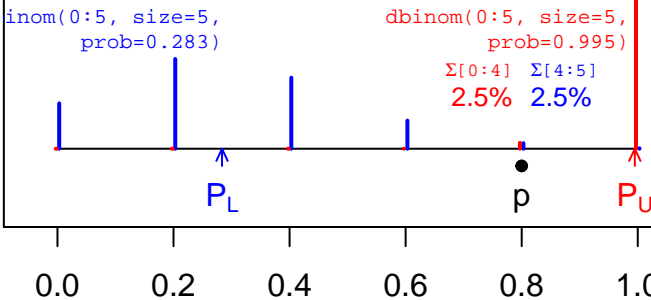
WILSON 1927. CI for proportion P, based on observed sample proportion p

Probable Inference (USUAL). Say we observe a certain proportion in a sample of n. We compute an interval using a statistical model (binomial or Gaussian) that uses (the statistic) p as the parameter for the sampling distribution.

It is common to say that the probability that the true proportion, P, lies below/above the 2.5/97.5%-ile of p (this is a sampling distribution) centered on p is 0.05:



Clopper-Pearson (1934) used 2 Binomial distributions



WILSON 1927 (continued...)

Strictly speaking, this statement is elliptical. Really the chance that P lies outside a specified range is either 0 or 1. It is the observed proportion p which has a greater or less chance of lying within a specified interval of P. If the observer was unlucky to have observed a rare event and to have based his inference thereon, he may be fairly vain of the mark.

Probable Inference (IMPROVED). A better way is to reason:
 There is some [true] P. Consider 2 scenarios:
 p --- p ('p is an under-estimate'):
 p landed at the 2.5%-ile of this sampling distribution (Distrn):
 $p = qDistrn(0.025, prob = P.Upper)$
 p landed at the 97.5%-ile of this sampling distribution (Distrn):
 $p = qDistrn(0.975, prob = P.Lower)$
 solve for P.Upper
 Wilson used 2 Gaussian sampling distributions
 P_L 16/20 P_U

Clopper-Pearson (1934) used 2 Binomial distributions

