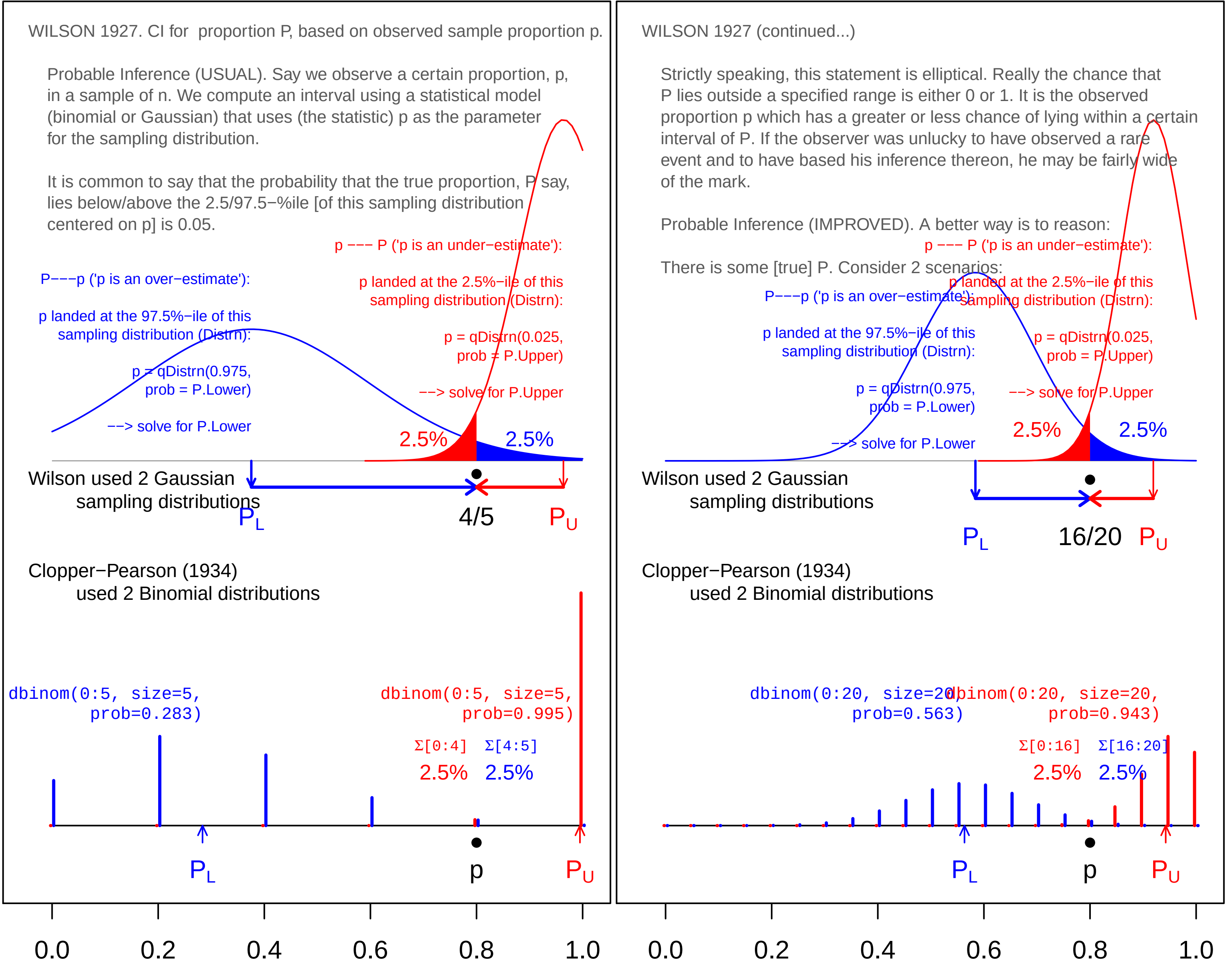

WILSON 1927. CI for proportion P, based on observed sample proportion p.
Probable Inference (USUAL). Say we observe a certain proportion, p, in a sample of n. We compute an interval using a statistical model (binomial or Gaussian) that uses (the statistic) p as the parameter for the sampling distribution.
It is common to say that the probability that the true proportion, P say, lies below/above the 2.5/97.5−%ile [of this sampling distribution centered on p] is 0.05.
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)
−−> solve for P.Upper
P−−−p ('p is an over−estimate'):
p landed at the 97.5%−ile of this sampling distribution (Distrn):
p = qDistrn(0.975, prob = P.Lower)
−−> solve for P.Lower
2.5%
2.5%
Wilson used 2 Gaussian sampling distributions
P_L
4/5
P_U
Clopper−Pearson (1934) used 2 Binomial distributions
dbinom(0:5, size=5, prob=0.283)
dbinom(0:5, size=5, prob=0.995)
Σ[0:4]
Σ[4:5]
2.5%
2.5%
P_L
p
P_U
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 certain 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 wide 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)
−−> solve for P.Upper
P−−−p ('p is an over−estimate'):
p landed at the 97.5%−ile of this sampling distribution (Distrn):
p = qDistrn(0.975, prob = P.Lower)
−−> solve for P.Lower
2.5%
2.5%
Wilson used 2 Gaussian sampling distributions
P_L
16/20
P_U
Clopper−Pearson (1934) used 2 Binomial distributions
dbinom(0:20, size=20, prob=0.563)
dbinom(0:20, size=20, prob=0.943)
Σ[0:16]
Σ[16:20]
2.5%
2.5%
P_L
p
P_U
0.0
0.2
0.4
0.6
0.8
1.0
0.0
0.2
0.4
0.6
0.8
1.0