Bayesian Methods for Sample Size Calculation

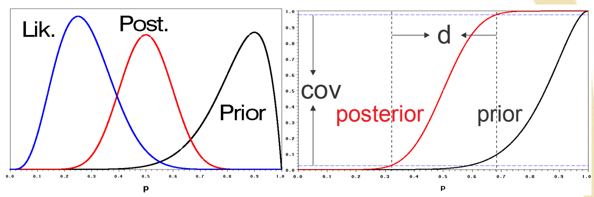
Hans Peter Stüger, Klemens Fuchs, Peter Much
AGES - Austrian Agency for Health & Food Safety

Abstract: Sample sizes for estimating proportions of agents present or antimicrobial resistance in study populations can be reduced with a Bayesian approach. Historical information out of former monitoring results was used to calculate the necessary sample size.

What`s the problem? The proportion p of resistant isolates against a set of antimicrobials has to estimated for several bacterial- and animals species by an annual monitoring program. To reduce costs by lower sample sizes results from previous years should be used.

1. How to use prior knowledge and historical information?
Bayesian approach: prior knowledge + sample data (likelihood) => posterior function

1a.



Credibility intervals with pre-specified coverage (cov) or length (d) can be calculated

1b.
Prior information: *Campylobacter* spp. in poultry
resistant isolates 2004: $x_{04}=22$ $(n_{04}=151)$
2005: $x_{05}=33$ $(n_{05}=205)$

How to weight this information?
=> older information is less weighted
2004: 25% 2005: 50%

2a. How to build in the weighted information?
=> conjugate prior function: Beta (a_0, b_0)
Beta coefficients are equivalent to a (virtual) sample of size $\mathbf{n_0 = a_0 + b_0 - 2}$ with $\mathbf{a_0 - 1}$ positive cases and $\mathbf{b_0 - 1}$ failures.

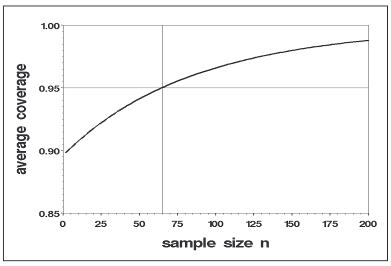
2b.
$a_0 = 0.25 \cdot x_{04} + 0.5 \cdot x_{05} + 1$
$b_0 = 0.25 \cdot (n_{04} - x_{04}) + 0.5 \cdot \cdot (n_{05} - x_{05}) + 1$

=> Beta (23 ; 119) as prior function

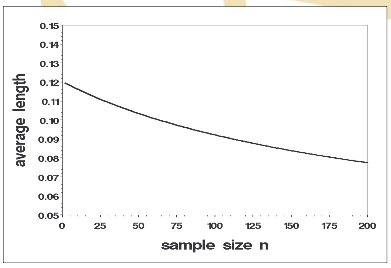
3. How to determine optimal sample size?
How big has n to be to get an
a) **average coverage AC** for a pre-specified interval length $\mathbf{d_0}$, or
b) **average length AL** for a pre-specified coverage $\mathbf{cov_0}$?

$AC(n) = \Sigma cov(x) \cdot m(x)$ .....x=0,...,n
$AL(n) = \Sigma d(x) \cdot m(x)$
m(x).......marginal probability

4. Average coverage for d_0=0.1



Average length for cov_0=0.95



5. Results: $\mathbf{n_{opt}=65}$

Classical approach: confidence interval (CI) of 0.15 ± 0.05 (95%) requires **n= 196**

A minimum sample size of n=45 was defined for a CI of 0.5 ± 0.15 (95%) out of sample alone.

Conclusions: Bayesian approach
- **allows incorporation of historical knowledge,**
- **reduces optimal sample sizes (and costs!) with equal precision and confidence level**

Contact: **Dr. Hans Peter Stüger,**
Austrian Agency for Health and Food Safety, Institute of Biostatistics, Beethovenstraße 8, 8010 Graz, Austria
hans-peter.stueger@ages.at, Phone: 0043/50555-64101