Fuzzy Set Representation of a Prior Distribution

Glen Meeden
University of Minnesota

http://www.stat.umn.edu/~glen/talks

Ordinary Confidence Intervals

OK for continuous data, but a really bad idea for discrete data.

Why?

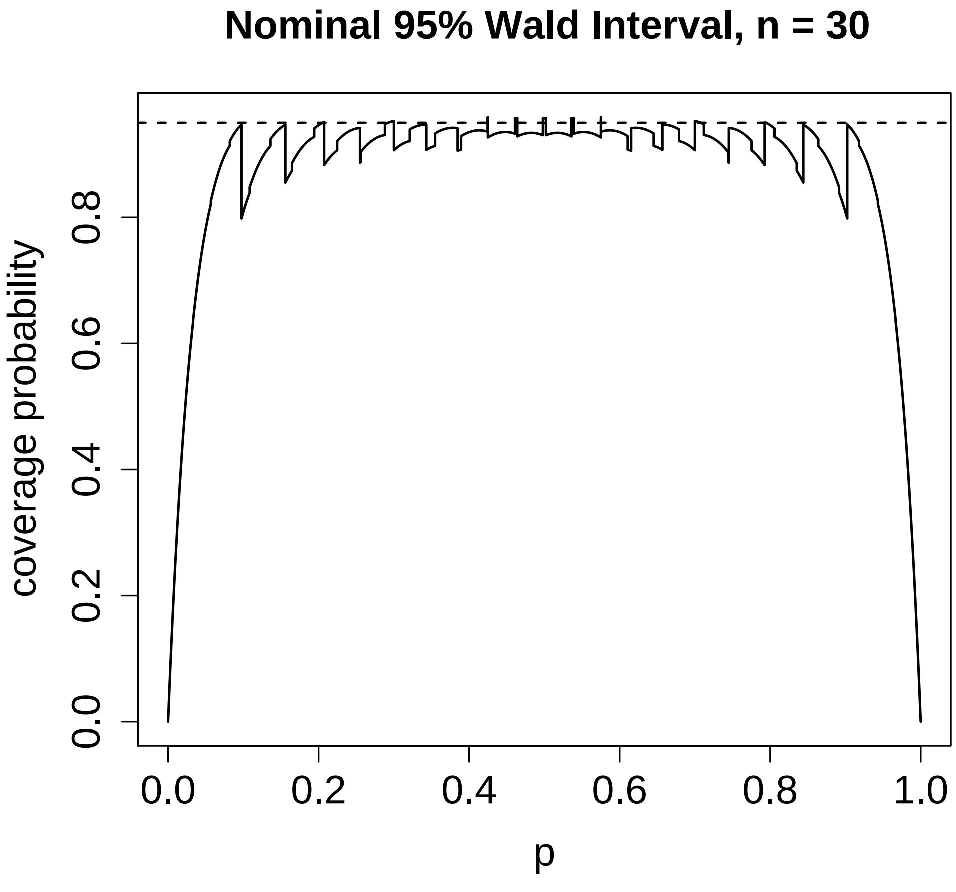
Coverage Probability

$$\begin{aligned} \gamma(\theta) &= \text{pr}_\theta\{l(X) < \theta < u(X)\} \\ &= \sum_{x \in S} I_{(l(x), u(x))}(\theta) \cdot f_\theta(x) \end{aligned}$$

As θ moves across the boundary of a possible confidence interval $(l(x), u(x))$, the coverage probability jumps by $f_\theta(x)$.

Ideally, γ is a constant function equal to the nominal confidence coefficient. But that's not possible.

Usual Binomial Interval



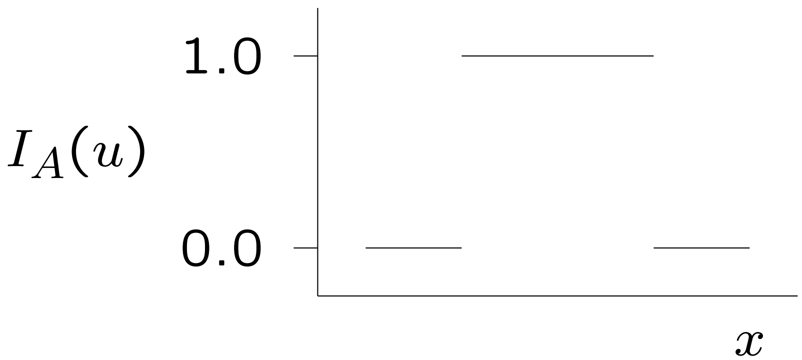
Performance of usual (Wald) interval

$$\hat{p} \pm 1.96\sqrt{\frac{\hat{p}(1-\hat{p})}{n}}$$

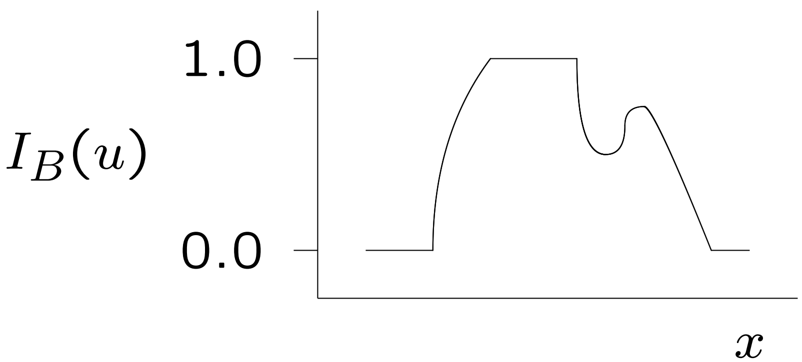
for Binomial$(30, p)$. Dotted line is nominal level (0.95).

Fuzzy Sets

Indicator function I_A of ordinary set A



Membership function I_B of fuzzy set B



Interpreting a fuzzy membership function

The value $I_B(u)$ is the degree of membership of the point u in the fuzzy set B.

Ordinary sets are special case of fuzzy sets called *crisp sets*.

A non-probabilistic measure of uncertainty.

Think partial credit! Geyer and Meeden (*Statist. Sci.*, 2005)

Randomized Tests

Randomized test defined by *critical function* $\phi(x, \alpha, \theta)$.

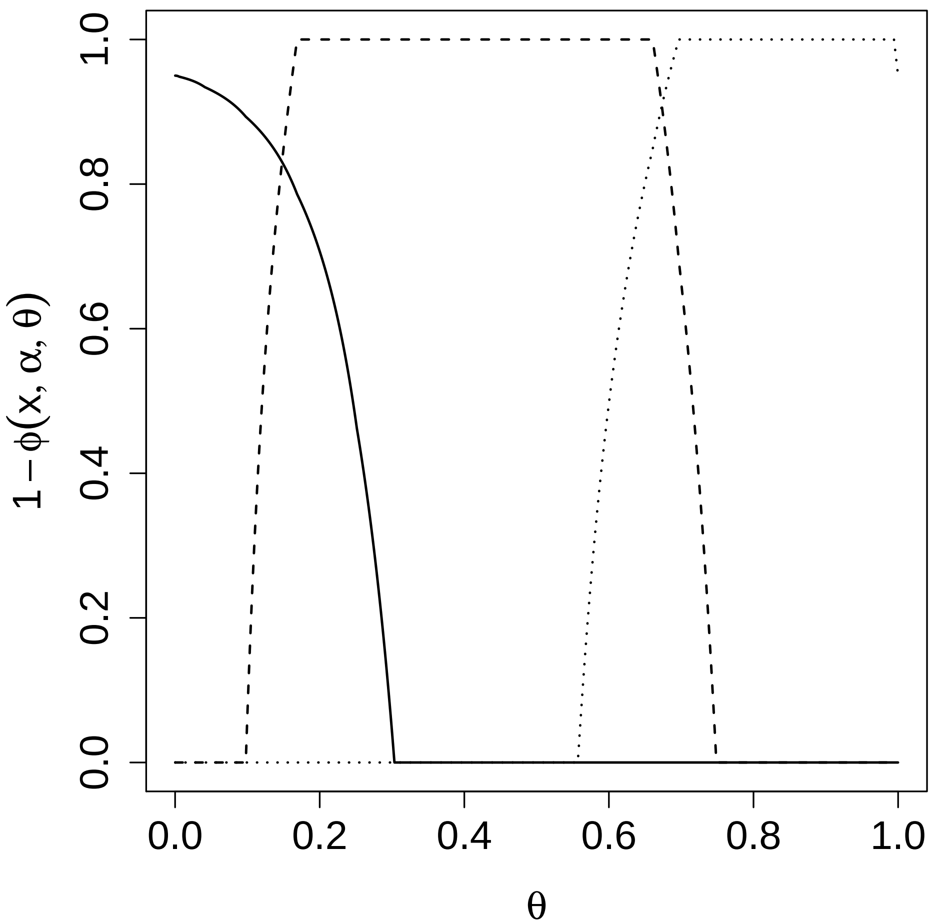
- observed data x.
- significance level α
- null hypothesis $H_0 : \theta = \theta_0$

Decision is randomized: reject H_0 with probability $\phi(x, \alpha, \theta_0)$.

Since probabilities are between zero and one, so is $\phi(x, \alpha, \theta)$.

Classical uniformly most powerful (UMP) and UMP unbiased (UMPU) tests are randomized when data are discrete.

Binomial Example



Sample size $n = 10$, fuzzy confidence interval associated with UMPU test, confidence level $1 - \alpha = 0.95$.

Data
$x = 0$ (solid curve)
$x = 4$ (dashed curve)
$x = 9$ (dotted curve)

Fuzzy and Bayes

Bayesian credible intervals have the same problem as CI's for discrete data.

Subjective Bayesians will not care since frequentist properties of a procedure are not important to them.

Typically Bayesians are not interested in the fuzzy notion of uncertainty. An exception are those interested in imprecise probability. See Gert de Cooman (*Fuzzy sets and systems*, 2005)

How can a Bayesian change their prior or posterior into a fuzzy set membership function which gives a good representation of their uncertainty?

A Bayesian decision problem

$\theta \in \Theta$ an interval of real numbers.

$\pi(\theta)$ a continuous prior density

$A \in \mathcal{A}$, the class of measurable membership functions on Θ.

The real numbers $0 < r_1 < r_2$ define the loss function

$$L(A,\theta) = \frac{1}{r_2 - r_1}\{1 - I_A(\theta)\} + \int_\theta \left(\frac{I_A(\theta)^2}{2} + \frac{r_1}{r_2 - r_1} I_A(\theta) \right) d\theta$$

The solution

The fuzzy set membership A which satisfies

$$\int_{\Theta} L(A, \theta)\pi(\theta)\, d\theta = \inf_{A' \in \mathcal{A}} \int_{\Theta} L(A', \theta)\pi(\theta)\, d\theta$$

is given by

$$I_A(\theta) = \begin{cases} 0 & \text{for} \quad 0 \le \pi(\theta) < r_1 \\ (\pi(\theta) - r_1)/(r_2 - r_1) & \text{for} \quad r_1 \le \pi(\theta) \le r_2 \\ 1 & \text{for} \quad \pi(\theta) > r_2 \end{cases}$$

A fuzzy membership function to a family of priors

Let ℓ = length of $\Theta < \infty$. Let A be a fuzzy membership function. If $r_1 < 1/\ell$ then there exists a unique $r_2 > r_1$ such that

$$\pi_{A,r_1}(\theta) = (r_2 - r_1)I_A(\theta) + r_1 \quad \text{for } \theta \in \Theta$$

is a prior density and for which the solution to our decision problem is $I_A(\theta)$. Furthermore a prior density π will have $I_A(\theta)$ as the solution to our decision problem if and only if it satisfies

$$\begin{aligned} \pi(\theta) &\geq r_2 \quad \text{when} \quad I_A(\theta) \geq 1 \\ \pi(\theta) &= I_A(\theta) \quad \text{when} \quad 0 < I_A(\theta) < 1 \\ \pi(\theta) &\leq r_1 \quad \text{when} \quad I_A(\theta) = 0 \end{aligned}$$

Final comment

We have seen how to convert a prior or posterior density into a fuzzy membership function. Formally this increases a Bayesian's uncertainty but in some cases it could be a good thing. Why?

For a Bayesian the posterior distribution summarizes their information about the parameter given the data. Bayesian credible intervals are a way to summarize the information in the posterior. These intervals have little formal justification in the Bayesian paradigm and seem to be popular because they mimic frequentist CI's. For discrete problems it may more useful to convert the posterior to a fuzzy membership function.

Perhaps frequentist and Bayesians should be more interested in fuzzy membership functions as a way to represent uncertainty about a parameter value.