Bayesian Clinical Trials

Sample Size for Binomial Proportions

Sample size based on interval widths

What sample size is needed to provide sufficient information to specify the true difference between two proportions π_1 and π_2 to within a total 100(1-α) interval width of given percentage points (say w)?

$$n = 4Z^2_{1-\alpha/2}\frac{\left[\pi_1\left(1-\pi_1\right)+\pi_2\left(1-\pi_2\right)\right]}{w^2}$$

This formula requires point estimates of π_1 and π_2 while a better summary of the available information is a distribution over a range of values.

Bayesian approach

Let θ in Θ the parameter of interest. We recall that the posterior distribution is

$$f(\theta|x) = \frac{L(x|\theta)\, g(\theta)}{\int_\Theta L(x|\theta)\, g(\theta)\, d\theta}$$

and depends on the data x, which is of course unknown at the planning stages of the experiment. The predictive distribution is

$$f(x) = \int_\Theta L(x|\theta)\, g(\theta)\, d\theta$$

Typically, we wish a highest posterior density (HPD) or other posterior credible interval of length l that covers θ with probability $(1-\alpha)$. HPD are optimal in the sense that they lead to the smallest sample sizes for any given coverage.

Criteria based on HPD (Joseph et al, 1997)

- ACC: controls the coverage rate of fixed length credible intervals over the predictive distribution of the data
- ALC: controls the length of credible intervals with a fixed coverage rate over the predictive distribution of the data.
- WOC: guarantees that the desired coverage rate and interval length over all (or a subset of) possible datasets.

Average Coverage Criterion (ACC)

The ACC sample size is the smallest integer such that for a fixed nominal interval length l the expected coverage level is at least $1-\alpha$, where expectation is taken over the marginal distribution of x:

$$\int \left(\int_{a(x,n)}^{a(x,n)+l} f\left(\theta|x\right) d\theta \right) f\left(x\right) dx \geq 1-\alpha$$

where $a\left(x,n\right)$ is the lower limit of the HPD interval of length l for the posterior density.

Average Length Criterion (ALC)

The ALC sample size is the smallest integer such that for a fixed nominal coverage level $1 - \alpha$ the expected length is at most l, where expectation is taken over the marginal distribution of x. The ALC has a two-step formula. The first step is to find the HPD length $l'(x, n)$ that satisfies

$$\int_{a(x,n)}^{a(x,n)+l'(x,n)} f(\theta|x)\, d\theta = 1 - \alpha$$

Then choose the minimum n that satisfies:

$$\int l'(x, n)\, f(x)\, dx \leq l$$

Worst Outcome Criterion (WOC)

The WOC sample size is determined by the smallest integer n such that

$$\inf \left(\int_{a(x,n)}^{a(x,n)+l} f\,(\theta|x)\, d\theta \right) \geq 1 - \alpha$$

Comparison between ALC, ACC and WOC (Cao et al, 2009)

- Because most researchers tend to report the length of the confidence interval of a fixed coverage, instead of the coverage of the interval of a fixed length, the ALC is more conventional than the ACC.
- WOC provides a conservative sample size and can give more assurance than the "average" assurances provided by the ACC and ALC criteria.
- When a posterior density has the common density shape of concave flat tails and convex steep center, the nominal coverage $1-\alpha$ will determine the relative difference between the ALC sample size and the ACC sample size. A "small" α implies that ACC needs a larger sample size than the ALC. It is the opposite for "bigger" α.

The difference of two binomial proportions

Let x_1 and x_2 be the total number of successes out of n_1 and n_2 trials from independent binomial experiments with parameters π_1 and π_2, respectively. We can choose two independent beta priors

$$g(\pi_1) = Beta(c_1, d_1) = B(c_1, d_1)^{-1} \pi_1^{c_1-1} (1-\pi_1)^{d_1-1}$$
$$g(\pi_2) = Beta(c_2, d_2) = B(c_2, d_2)^{-1} \pi_2^{c_2-1} (1-\pi_2)^{d_2-1}$$

for π_1 and π_2, respectively - each conjugate to the binomial likelihood - so the posterior distribution is the product of two Beta distributions:

$$f(\pi_1, \pi_2 | x_1, x_2, n_1, n_2) =$$
$$= \frac{\prod_{i=1}^{2} \pi_i^{c_i+x_i-1} (1-\pi_i)^{d_i+n_i-x_i-1}}{B(c_1+x_1, d_1+n_1-x_1) \, B(c_2+x_2, d_2+n_2-x_2)}$$

and the predictive distribution is the product of two betabinomial distributions:

$$p(x_1, x_2) =$$
$$= \prod_{i=1}^{2} \left[B(c_i + x_i, d_i + n_i - x_i) B(c_i, d_i)\right]^{-1} \pi_i^{c_i + x_i - 1} (1 - \pi_i)^{d_i + n_i - x_i - 1}$$

We make a change of variable as $\theta = \pi_2 - \pi_1$ and the posterior becomes:

$$f(\pi_1, \theta | x_1, x_2, n_1, n_2) \propto$$
$$\propto \pi_1^{c_1 + x_1 - 1} (1 - \pi_1)^{d_1 + n_1 - x_1 - 1} (\pi_1 - \theta)^{c_2 + x_2 - 1} (1 - \pi_1 + \theta)^{d_2 + n_2 - x_2 - 1}$$

It follows that the marginal posterior distribution of θ is

$$f(\theta | x_1, x_2, n_1, n_2) = \int_{\max 0, \theta}^{\min \theta + 1, 1} f(\pi_1, \theta | x_1, x_2, n_1, n_2)\, d\pi_1$$

- ACC sample size for θ is the minimum n:

$$\sum_{x_1=0}^{n}\sum_{x_2=0}^{n}\int_{a(x_1,x_2)}^{a(x_1,x_2)+l} f\left(\theta|x_1,x_2,n\right)d\theta \;\; p\left(x_1,x_2\right)\geq 1-\alpha$$

- ALC sample size for θ is the minimum n:

$$\sum_{x_1=0}^{n}\sum_{x_2=0}^{n} l'\left(x_1,x_2\right)p\left(x_1,x_2\right)\leq l$$

where $l'\left(x_1,x_2\right)$ is the HPD length that satisfies:

$$\int_{a(x_1,x_2)}^{a(x_1,x_2)+l'(x_1,x_2)} f\left(\theta|x_1,x_2,n\right)d\theta = 1-\alpha$$

- WOC sample size for θ is the minimum n:

$$\int_{a\left(x_1^*,x_2^*\right)}^{a\left(x_1^*,x_2^*\right)+l} f\left(\theta|x_1^*,x_2^*,n\right)d\theta \geq 1-\alpha$$

where x_1^* and x_2^* are the numbers of successes that maximize the length of the HPD interval.

References

- Joseph L, du Berger R, Bélisle P. Bayesian and mixed Bayesian/likelihood criteria for sample size determination. Stat Med 1997; 16(7): 769-781.
- Cao J, Lee JJ, Alber S. Comparison of Bayesian sample size criteria: ACC, ALC, and WOC. J Stat Plan Inference 2009; 139(12):4111-4122.