Decision Tools for Adaptive Designs

Werner Brannath

VO ”Sequential and Adaptive Designs”,
University Bremen

Conditional Power

Conditional Power

Definition of Conditional Power (CP):

Conditional probability to reject the null given the stage 1 data and a parameter value θ belonging to the alternative.

CP with Conditional error function approach:

$$q \leq A(\text{stage 1 data}) \quad \Longleftrightarrow \quad \Phi^{-1}(q) \leq \Phi^{-1}\big(A(\text{stage 1 data})\big)$$

Typically

$$\Phi^{-1}(q) \sim N\big(-\theta\sqrt{I_2}, 1\big)$$

where I_2 is the (incremental) information of stage 2. Hence,

$$CP_\theta = P_\theta[\text{reject}|\text{stage 1 data}] = \Phi\left(\Phi^{-1}\big(A(\text{stage 1 data})\big) + \theta\sqrt{I_2}\right)$$

$$= 1 - \Phi\left(\Phi^{-1}\big(1 - A(\text{stage 1 data})\big) - \theta\sqrt{I_2}\right)$$

Explicit formula

With a two-arm z-test: $I_2 = n_2/(4\sigma^2)$; n_2 second stage sample size.

Fisher's product test:

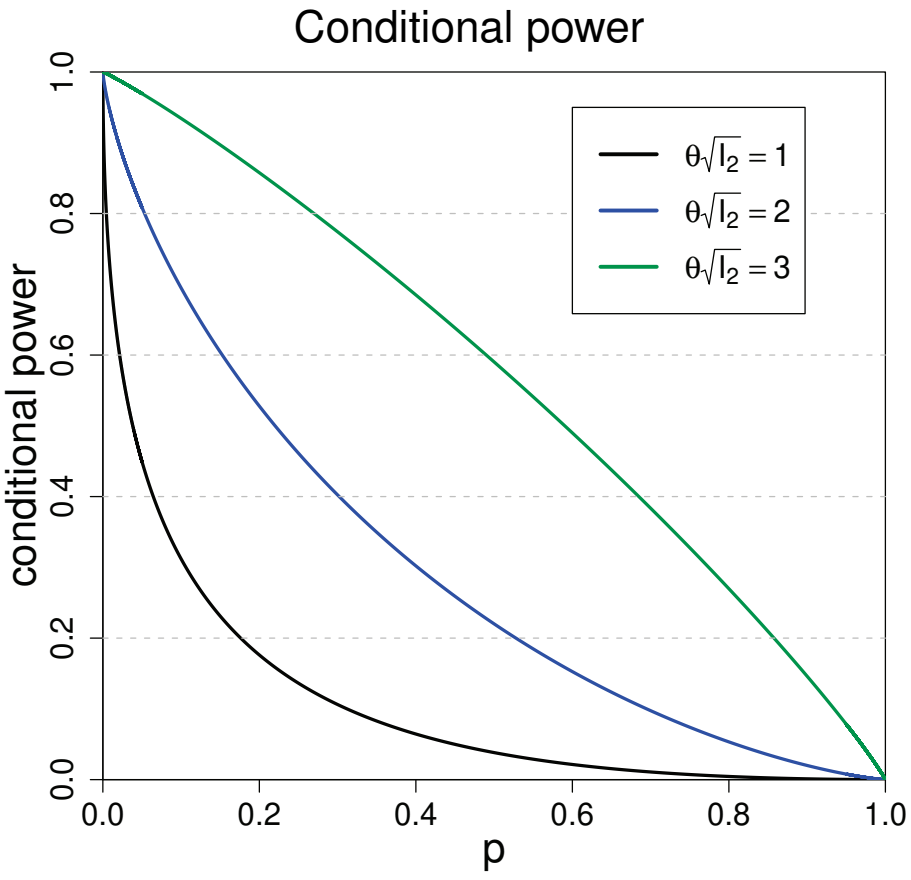
$$CP_\theta = \Phi\left(\Phi^{-1}\left(\max\left(\frac{c}{p}, 1\right)\right) + \theta\sqrt{I_2}\right)$$

Inverse normal method:

$$CP_\theta = \Phi\left(\frac{w_1 Z_1 - u_2}{w_2} + \theta\sqrt{I_2}\right)$$

Same conditional power for GSD, then with $w_1 = \sqrt{n_1/(n_1 + n_2)}$ (planned sample sizes).

Conditional Power – Properties



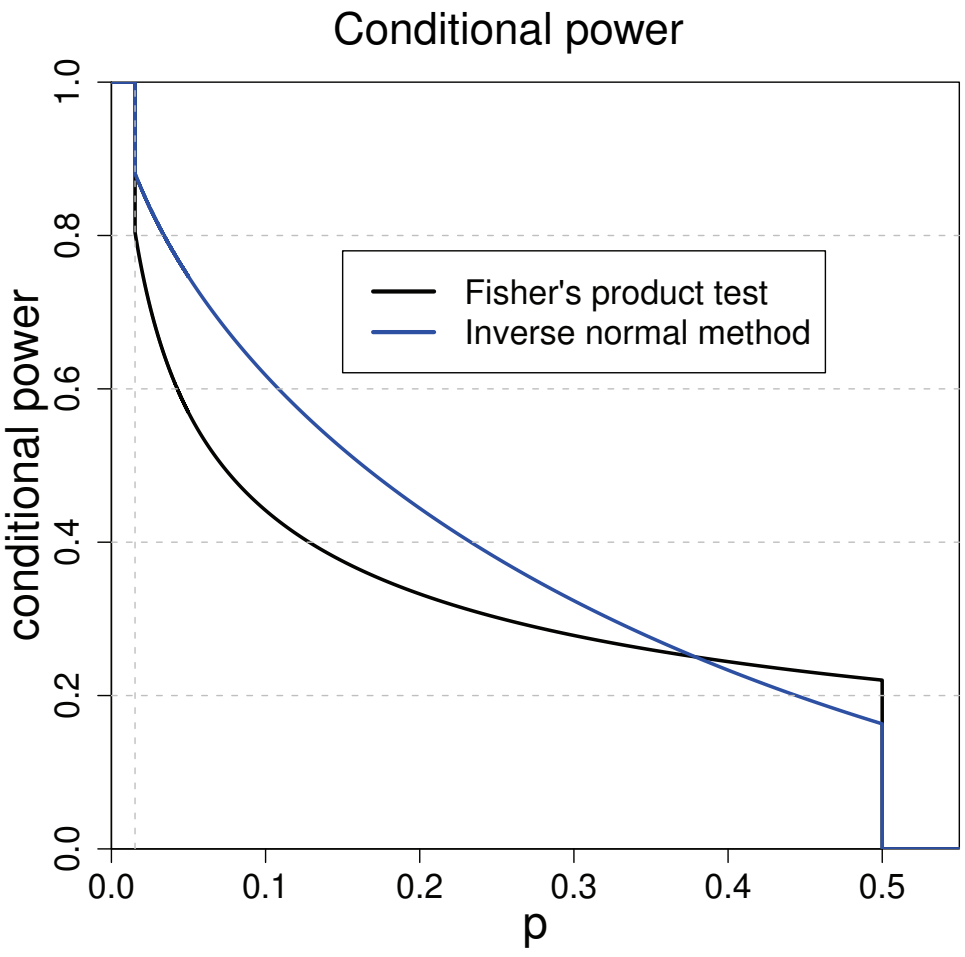
The conditional power

- increases with increasing θ;
- increases with increasing I_2 (second stage sample size);
- decreases with increasing p

Example:

- Inverse normal method
- $\alpha = 0.025$
- $\alpha_0 = 1, \alpha_1 = 0$

Conditional Power – Examples



- θ_t, θ_c . . . mean responses under t and c, resp.
- $H_0 : \theta_t \leq \theta_c$ vs. $H_1 : \theta_t > \theta_c$
- $\alpha = 0.025$
- $n_1 = 100$
- $\alpha_0 = 0.5$
- $\alpha_1 = 0.015$
- $\theta_0 = 0.25$

Early stopping with CP

- It has been suggest (already for GSD) to stop the trial for futility if $CP_\theta(\text{stage 1 data})$ is small, i.e. below some cp_0, for the planning alternative θ.
- There is one-to-one correspondence between α_0 and cp_0:

$$cp_0 = \Phi\left(\Phi^{-1}\left(A(\alpha_0)\right) + \theta\sqrt{I_2}\right)$$

$$\alpha_0 = A^{-1}\left(\Phi\left(\Phi^{-1}(cp_0)\right) - \theta\sqrt{I_2}\right)$$

- E.g. $\alpha = 0.05$, $\alpha_0 = 0.5$, $\alpha_1 = 0.0233$ and power is 90%;

  Inverse normal combination test ($w_1 = \sqrt{0.5}$): $cp_0 = 0.33$;

  Fisher's product test: $cp_0 = 0.48$ (see Figure 7.1 in WaBr16).

Sample size adaptations

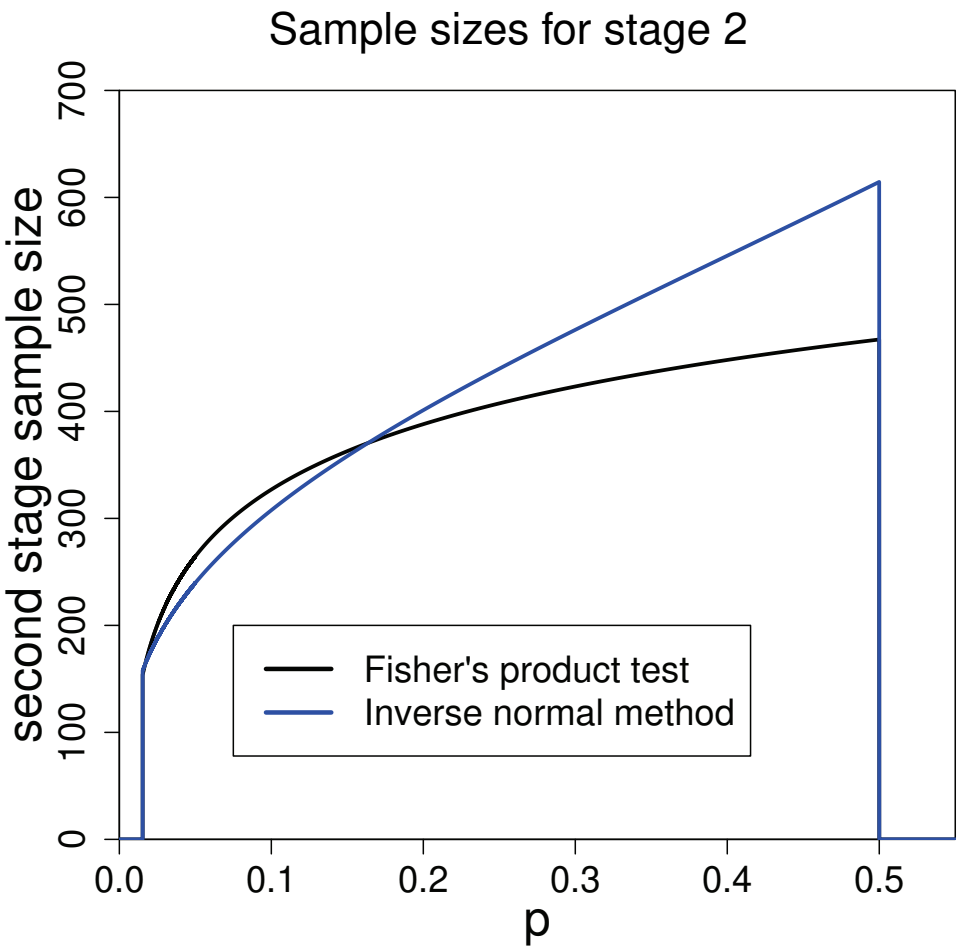
Sample size adaptations based on conditional power

- Given interim data either stop the trial or choose the stage 2 sample size such that the conditional power is at least π, e.g., $\pi = 0.8$.
- *Rational:* If we continue with the trial then we want to have a "good chance" ($\geq \pi$) to finally reject the null hypothesis.
- If we continue with stage 2 then we choose the second stage sample size

$$n_2 = \left\{ \Phi^{-1}(\pi) - \Phi^{-1}\left(A[\text{stage 1 data}]\right) \right\}^2 / (I_1 \cdot \theta^2)$$

where I_1 is the information per observation (e.g. $I_1 = 0.5/\sigma^2$ in a two-armed clinical trial).

Sample size adaptations based on conditional power



- θ_t, θ_c . . . mean responses under t and c, resp.
- $H_0 : \theta_t \leq \theta_c$ vs. $H_1 : \theta_t > \theta_c$
- $\alpha = 0.025$
- $n_1 = 100$
- $\alpha_0 = 0.5$, $\alpha_1 = 0.015$
- stage 2 sample size s.th.

$$CP_{\theta_0=0.25} = 0.9$$

- Overall power 90%,

  independent from c.e.f.

Which effect in the conditional power calculation?

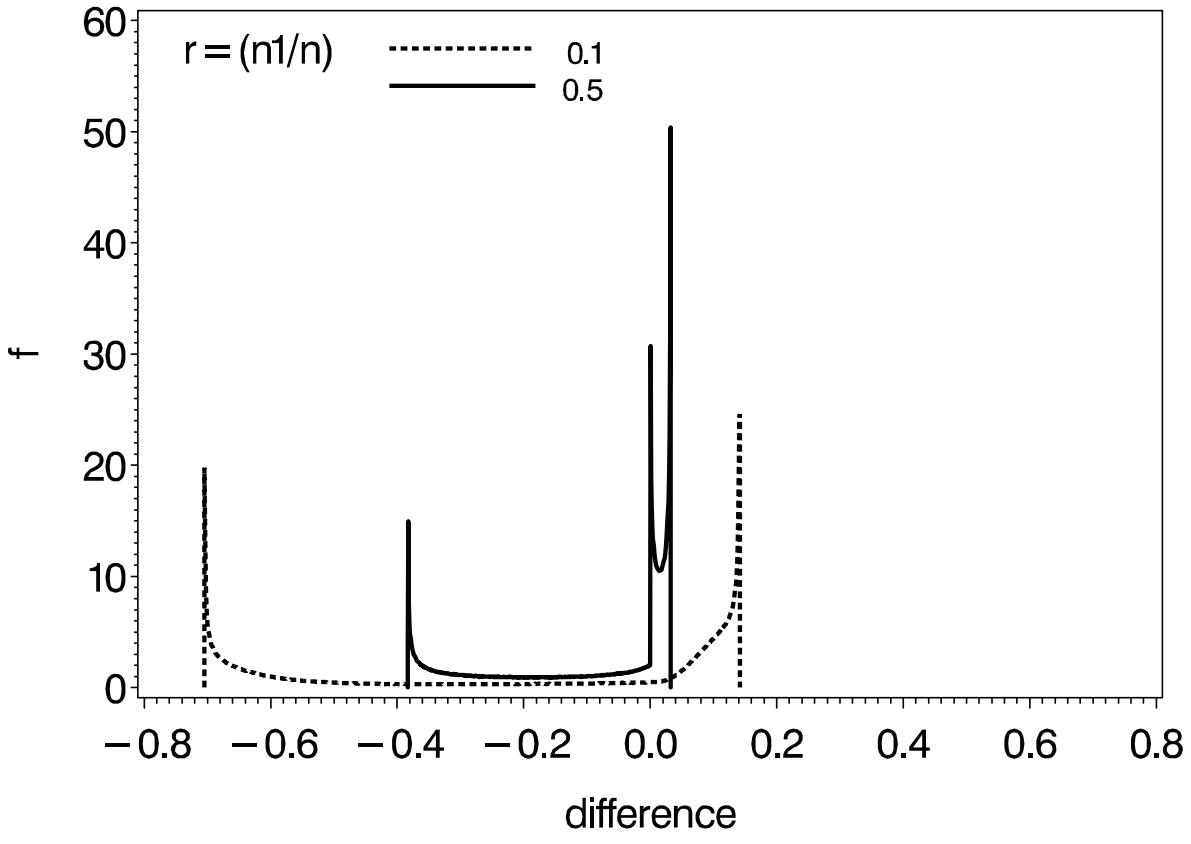
Several possibilities have been considered

- Using the effect size θ_0 initially used for planning the trial (θ_0 minimal relevant effect size).
- Using the interim estimate $\hat{\theta}_1$ of θ (if belonging to the alternative) in the hope to estimate the conditional power under true unknown effect.
- Using a weighted sum of the initial and estimated effect size or a posterior mean for some given prior distribution:

$$\tilde{\theta}_1 = \theta_0 \frac{I_0}{I_0 + I_1} + \hat{\theta}_1 \frac{I_1}{I_0 + I_1}$$

where θ_0 is the *prior mean* (initial guess), I_1 is the interim and I_0 the prior information on θ.

How good can we estimate the true conditional power? (Bauer and König 2006)



Density of difference between true and estimated CP:

- median is 0;
- density spreads to left much more than to the right
- relatively high chance for a substantial underestimation of true CP

How good can we estimate the true conditional power? – Conclusions

(Bauer and König 2006)

- Using the interim estimate for estimating the conditional power can be quite misleading and can lead to a severe underestimation of the true conditional power.
- One should not over-interpret the interim data.
- It seems better to use the minimal relevant alternative θ_0,

  and, if necessary, to adjusted it only **carefully** by what we have learned at stage 1.
- *Graphical tool:* Plot of CP_θ over a range of plausible θ values together with the likelihood (or prior density) for θ.

  (See Figure 7.4 (p. 185) in WaBr16 for an example.)

Clinical Trial Example

Clinical trial example - study plan (Zajicek et al., 2012)

- Phase III trial on 12 week trt with oral cannabis extract (CE) for the symptomatic relief of muscle stiffness and pain in adults with stable multiple sclerosis (MS).
- Primary endpoint was relief in muscle stiffness (y/n) from baseline to 12 weeks (in 11 point category rating scale, CRS).
- Pre-planned sample size, based on previous study data, was 200 patients per arm (400 in total).
- Unblinded interim analysis planned after first 200 patients.
- Inverse normal combination test with equal weights.
- Early rejection with O'Brien & Fleming boundary ($\alpha_1 = 0.0026$).
- Sample size adaptation based on conditional power.

Clinical trial example - interim results

- 101 pats in CE arm; 97 in placebo arm.
- Relief in muscle stiffness: 27 pats in CE (27% relief) and 12 in placebo arm (12.3% relief).
- First stage one-sided p-value was 0.0055 ($> \alpha_1 = 0.0026$)
- 250 pats already randomized
- Conditional power with 300 pats $\geq$ 90% under pre-planned and estimated sample size.
- Therefore decision (by iDMC) to reduce sample size from 400 to 300 pats in total.

Clinical trial example - final results

- Study stopped with 143 pats in CE arm; 134 in placebo arm.
- Overall relief rates: 0.294 in CE and 0.157 in placebo arm.
- Stage two rates: 35.7% in CE and 24.3% in placebo arm.
- Weighted z-score was 2.573 exceeding the critical boundary of 1.977.
- Difference statistical significant.

- *Heterogeneity(?):* Increase of relief rates from stage 1 to stage 2.

  Systematic or due to chance?

  Second stage sample size small; no clear indication for bias.

Bayesian Predictive Power

Bayesian statistics - basic idea

- In a Bayesian statistical analysis with start *a priori* (or *prior*) probability distribution for the parameter of interest, e.g. a normal density $\pi_0(\theta)$ with mean θ_0 and variance $\sigma_0 = 1/I_0$.
- $\pi_0(\theta)$ describes our prior believes about θ.
- We need to assume a model for the data X with distributions $f_\theta(x)$ for the data for given parameter values θ.
- Given the data we the calculate (using Bayes' rule) the conditional density of θ given the data x:

$$\pi_1(\theta|x) = \frac{\pi_0(\theta) f_\theta(x)}{\int_{-\infty}^{\infty} \pi_0(\theta') f_{\theta'}(x) d\theta'}$$

  We call $\pi_1(\theta|x)$ the *posterior* density of θ (*posterior* distribution).
- $\pi_1(\theta|x)$ describes our believes on θ after we have seen the data.

Bayesian statistics - use of posterior distribution

- We can e.g. report the *posterior* mean

$$\tilde{\theta}(x) := \int_{-\infty}^{\infty} \theta' \pi_1(\theta'|x)\, d\theta'$$

- If π_0 is the density of $N(\theta_0, I_0)$ and $X = \hat{\theta} \sim N(\theta, 1/I_1)$, then π_1 is a normal density with expectation

$$\tilde{\theta}_1 = \theta_0 \frac{I_0}{I_0 + I_1} + \hat{\theta}_1 \frac{I_1}{I_0 + I_1}$$

  (as on slide 11) and variance $1/(I_0 + I_1)$.

- 95% *credible interval* for θ is the interval $I = (\tilde{\theta}_1 - c, \tilde{\theta}_1 + c)$ with

$$\mathbf{P}(\theta \in I \,|\, X = x) = \int_I \pi_1(\theta'|x) d\theta' = 0.95$$

Bayesian Predictive Power

- We an use the interim prior density $\pi_1(\theta|\hat{\theta}_1)$ to calculate the chance for a final rejection:

$$PP_{\pi_0}(\hat{\theta}_1) = \mathbf{P}_{\pi_0}\left(\text{reject } H_0 \text{ at } 2^{nd} \text{ stage} \,\middle|\, \hat{\theta}_1\right) = \mathbf{P}_{\pi_0}\left(q \leq A(p) \,\middle|\, \hat{\theta}_1\right)$$

$$= \int_{-\infty}^{\infty} \underbrace{\mathbf{P}_{\theta}\left(q \leq A(p) \,\middle|\, \hat{\theta}_1\right)} \, \pi_1(\theta|\hat{\theta}_1)\, d\theta$$

$$= \int_{-\infty}^{\infty} CP_{\theta}(\hat{\theta}_1)\, \pi_1(\theta|\hat{\theta}_1)\, d\theta$$

- With the predictive power we account for the uncertainty in our estimate of the true effect.
- We ignore this uncertainty when using the prior mean $\tilde{\theta}_1$ for the (frequentist) conditional power, as indicated on slide 11.

Conditional power formula (reminder)

We we will later use $Z_2 = \Phi^{-1}(1-q) \sim N(\theta\sqrt{I_2}, 1)$ and the relation

$$CP_\theta(\hat{\theta}_1) = \mathbf{P}_\theta\left(Z_2 \geq \Phi^{-1}(1-A(p)) \,\middle|\, \hat{\theta}_1\right)$$

$$= 1 - \Phi\left(\Phi^{-1}\big(1 - A(\text{stage 1 data})\big) + \theta\sqrt{I_2}\right)$$

Plugging-in for θ the posterior mean $\tilde{\theta}_1$, ignores our uncertainty about θ expressed by the posterior variance $1/(I_0 + I_1)$.

Predictive distribution

- We can calculate $PP_{\pi_0}(\hat{\theta}_1)$ via the *predictive distribution* of

$$Z_2 = 1 - \Phi^{-1}(1-q)$$

  which is the (marg.) distribution of Z_2 with random $\theta \sim \pi_1(\cdot|\hat{\theta}_1)$.

- If $\pi_1(\cdot|\hat{\theta}_1)$ is the density of $N\big(\tilde{\theta}_1, 1/(I_0+I_1)\big)$ then

$$Z_2 = \underbrace{Z_2 - \theta\sqrt{I_2}}_{N(0,1)} + \theta\sqrt{I_2} \quad \sim \quad N(0,1) \ * \ N(\tilde{\theta}_1\sqrt{I_2}, I_2/(I_0+I_1))$$

$$= N\left(\tilde{\theta}_1\sqrt{I_2}\,,\ \frac{I_0+I_1+I_2}{I_0+I_1}\right)$$

Predictive power formula

- This gives the predictive power formula

$$PP_{\pi_0}(\hat{\theta}_1) = 1 - \Phi\left(\sqrt{\frac{I_1}{I_0 + I_1 + I_2}}\left(\Phi^{-1}(1 - A(p)) - \tilde{\theta}_1\sqrt{I_2}\right)\right)$$

- If $I_0 = 0$ (uninformative prior) then $\tilde{\theta}_1 = \hat{\theta}_1$ and predictive power becomes

$$PP_{\pi_0}(\hat{\theta}_1) = 1 - \Phi\left(\sqrt{\frac{I_1}{I_1 + I_2}}\left(\Phi^{-1}(1 - A(p)) - \hat{\theta}_1\sqrt{I_2}\right)\right)$$

- The comparison to conditional power formula on slide 22 shows:

  PP with uniformative prioir shrinks the CP towards 0.5