Pubh 8482: Sequential Analysis
Testing Normal Random Variables

Joseph S. Koopmeiners

Division of Biostatistics
University of Minnesota

Week 2

Comparing Normal Means with Known Variance

Recall our set-up from last time

- Let $X_1, X_2, \ldots, X_n$ be *i.i.d.* $N\left(\mu_x, \sigma^2\right)$
- Let $Y_1, Y_2, \ldots, Y_n$ be *i.i.d.* $N\left(\mu_y, \sigma^2\right)$
- σ^2 known

Null hypothesis

Consider a two-sided test of

$$H_0 : \mu_x = \mu_y \quad \textit{vs.} \quad H_a : \mu_x \neq \mu_y$$

Fixed-sample test

- Collect n subjects in each group
- Test null hypothesis using the following test statistic

$$Z_n = \frac{\bar{X}_n - \bar{Y}_n}{\sqrt{\frac{2*\sigma^2}{n}}}$$

- Under the null, $Z_n \sim N(0,1)$
- Reject if $|Z_n| > Z_{1-\alpha/2}$
- Results in type-1 error rate of α

Group Sequential Design

- Consider a group sequential design with K stopping times
- Define Z_k analogously to Z_n where Z_k uses all observations at the *kth* stopping time
- For now, assume that stopping time are equally spaced throughout the trial
  - That is, the sample size at the *kth* stopping time is $\frac{k}{K}n$

Definitions

Generalize definitions from previous lectures to K stopping times

- $\delta = \mu_x - \mu_y$
- $\hat{\delta}_k = \bar{X}_k - \bar{Y}_k$
- $I_k = \frac{k}{K} \frac{n}{2 * \sigma^2}$
- $Z_k = \hat{\delta}_k \sqrt{I_k}$

Sequential

For K stopping times, the sequence of test statistics $(Z_1, \ldots, Z_K)$ follows a multivariate normal distribution with

- $E[Z_k] = \delta\sqrt{I_k}$ for $k = 1, \ldots, K$.
- $Var[Z_k] = 1$ for $k = 1, \ldots, K$.
- $Cov[Z_{k_1}, Z_{k_2}] = \sqrt{Z_{k_1}/I_{k_2}}$ for $1 \leq k_1 \leq k_2 \leq K$.

A General Stopping Rule

- Define critical values c_k for $k = 1, \ldots, K$
- For $k = 1, \ldots, K - 1$
  - If $|Z_k| > c_k$, stop and reject H_0
  - otherwise, continue to group $k + 1$
- For $k = K$
  - If $|Z_K| > c_K$, stop and reject H_0
  - otherwise, stop and fail to reject H_0

Example

Let $K = 4$

- $|Z_1| < c_1$, continue to group 2
- $|Z_2| < c_2$, continue to group 3
- $|Z_3| > c_3$, stop and reject H_0

Example

Let $K = 4$

- $|Z_1| < c_1$, continue to group 2
- $|Z_2| < c_2$, continue to group 3
- $|Z_3| < c_3$, continue to group 4
- $|Z_4| < c_4$, stop and fail to reject H_0

Key Question

Key Question:

- How do we choose the critical values?

Pocock Bounds

- The simplest approach to finding critical values is to use the approach proposed by Pocock (1977)
- Identify a constant critical value that provides the correct overall type-I error rate
- That is
  - $c_k = C_{PK}(\alpha, K)$ for all $k = 1, \ldots, K$
  - Find $C_{PK}(\alpha, K)$ that provides the desired type-I error rate
  - Note that $C_{PK}(\alpha, K)$ is a function of α and K

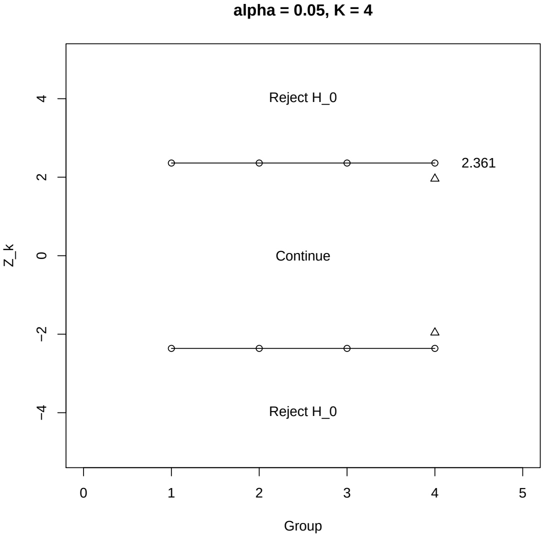
Pocock Bounds

- Formally, the Pocock test is as follows:
- For $k = 1, \ldots, K - 1$
  - if $|Z_k| > C_{PK}(\alpha, K)$, stop and reject H_0
  - otherwise, continue to group $k + 1$
- For $k = K$
  - if $|Z_K| > c_{PK}(\alpha, K)$, stop and reject H_0
  - otherwise, stop and fail to reject H_0

Pocock Bounds: Example

- $\alpha = 0.05$ and $K = 4$: $C_{PK}(\alpha, K) = 2.361$
  - $c_1 = 2.361$
  - $c_2 = 2.361$
  - $c_3 = 2.361$
  - $c_4 = 2.361$

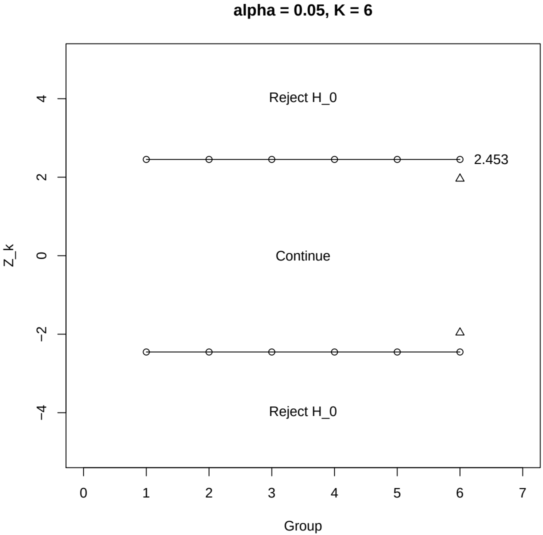
Pocock Bounds: Examples



Pocock Bounds: Example

- $\alpha = 0.05$ and $K = 6$: $C_{PK}(\alpha, K) = 2.453$
  - $c_1 = 2.361$
  - $c_2 = 2.361$
  - $c_3 = 2.361$
  - $c_4 = 2.361$
  - $c_5 = 2.361$
  - $c_6 = 2.361$

Pocock Bounds: Examples



$C_{PK}(\alpha, K)$ as a function of α and K

- $C_{PK}(\alpha, K)$ increases as K increases
- $C_{PK}(\alpha, K)$ increases as α decreases

Pocock: Advantages and Disadvantages

- Advantages
  - Simple!
  - Aggressive with regards to stopping early and, therefore, has small expected sample size
- Disadvantages
  - Substantial reduction in power requires and relatively large increase in maximum sample size

O'Brien-Fleming Bounds

- A second approach, and possibly the most popular, was proposed by O'Brien and Fleming in 1979
- O'Brien-Fleming Bounds using a very large critical value early in the study and use progressively smaller critical value as the study progresses
- Specifically:
  - $c_k = C_{OF}(\alpha, K)\sqrt{K/k}$ for all $k = 1, \ldots, K$
  - Find $C_{OF}(\alpha, K)$ that provides the desired type-I error rate

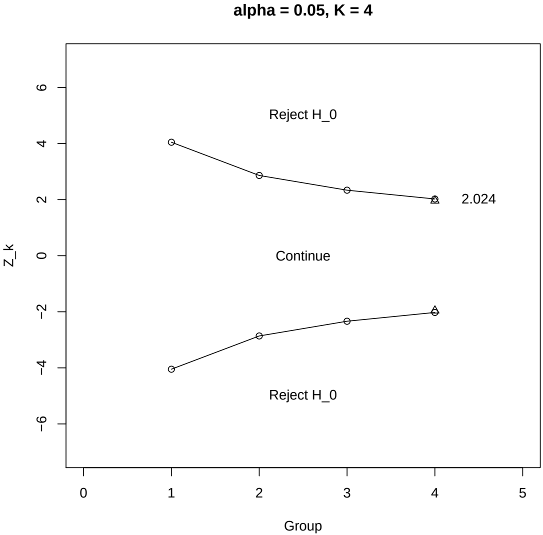
O'Brien-Fleming Bounds

- Formally, the O'Brien-Fleming test is as follows:
- For $k = 1, \ldots, K - 1$
  - if $|Z_k| > C_{OF}(\alpha, K)\sqrt{K/k}$, stop and reject H_0
  - otherwise, continue to group $k + 1$
- For $k = K$
  - if $|Z_K| > c_{OF}(\alpha, K)$, stop and reject H_0
  - otherwise, stop and fail to reject H_0

O'Brien-Fleming Boundaries: Example

- $\alpha = 0.05$ and $K = 4$: $C_{OF}(\alpha, K) = 2.024$
  - $c_1 = 4.048$
  - $c_2 = 2.862$
  - $c_3 = 2.337$
  - $c_4 = 2.024$

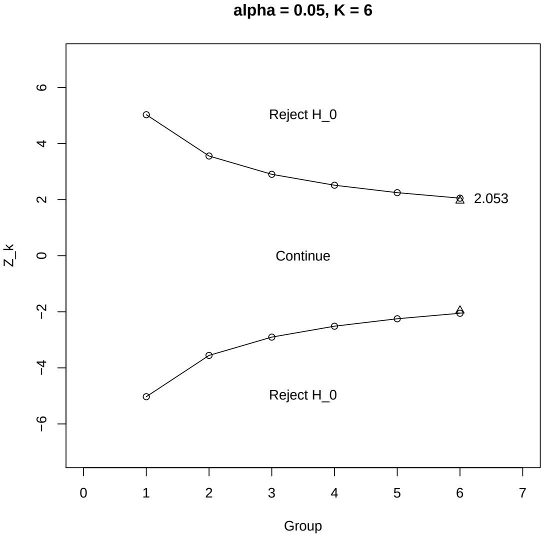
O'Brien-Fleming Boundaries: Example



O'Brien-Fleming Boundaries: Example

- $\alpha = 0.05$ and $K = 6$: $C_{OF}(\alpha, K) = 2.053$
  - $c_1 = 5.029$
  - $c_2 = 3.556$
  - $c_3 = 2.903$
  - $c_4 = 2.514$
  - $c_5 = 2.249$
  - $c_6 = 2.053$

O’Brien-Fleming Boundaries: Example



O'Brien Fleming: Advantages and Disadvantages

- Advantages
  - Final critical value is close to critical value for fixed-sample design
  - More powerful than Pocock and therefore requires a smaller maximum sample size
- Disadvantages
  - Less likely to stop early than Pocock boundaries and, therefore, has larger expected sample size

Wang and Tsiatis Family of Tests

- Wang and Tsiatis (1987) proposed a family of sequential tests indexed by a parameter Δ
- Specifically:
  - $c_k = C_{WT}(\alpha, K, \Delta)\,(k/K)^{\Delta - 0.5}$ for all $k = 1, \ldots, K$
  - $0 \leq \Delta \leq 0.5$
  - Find $C_{WT}(\alpha, K, \Delta)$ that provides the desired type-I error rate

Wang and Tsiatis Family of Tests

- Formally, the Wang and Tsiatis test is as follows:
- For $k = 1, \ldots, K - 1$
  - if $|Z_k| > C_{WT}(\alpha, K, \Delta)(k/K)^{\Delta - 0.5}$, stop and reject H_0
  - otherwise, continue to group $k + 1$
- For $k = K$
  - if $|Z_K| > c_{WT}(\alpha, K, \Delta)$, stop and reject H_0
  - otherwise, stop and fail to reject H_0

Relationship with Pocock and O'Brien-Fleming Tests

- The Pocock and O'Brien-Fleming tests are special cases of the Wang and Tsiatis Test
  - $\Delta = 0$ produces O'Brien-Fleming boundaries
  - $\Delta = 0.5$ produces Pocock Boundaries
  - $0 < \Delta < 0.5$ produces intermediate shapes

Wang and Tsiatis Boundaries

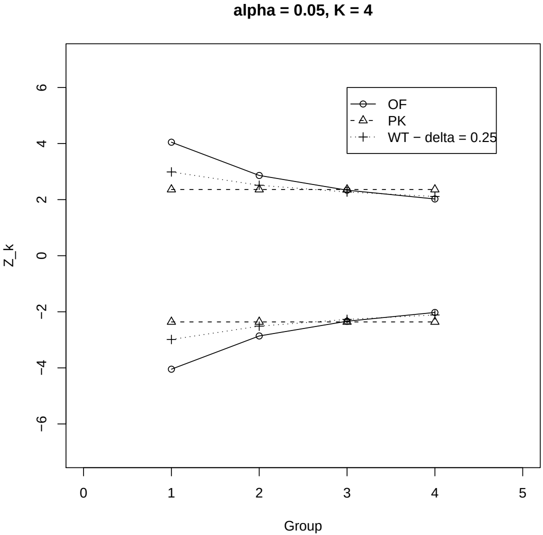
Examples

- $C_{WT}(\alpha, K, \Delta)$ is a function of α, K and Δ
- $C_{WT}(\alpha, K, \Delta)$ increases as Δ increases
  - $\alpha = 0.05$, $K = 6$ and $\Delta = 0.1$: $C_{WT}(\alpha, K, \Delta) = 2.083$
  - $\alpha = 0.05$, $K = 6$ and $\Delta = 0.25$: $C_{WT}(\alpha, K, \Delta) = 2.154$
  - $\alpha = 0.05$, $K = 6$ and $\Delta = 0.4$: $C_{WT}(\alpha, K, \Delta) = 2.292$
- $C_{WT}(\alpha, K, \Delta)$ increases as α and K

Wang and Tsiatis Boundaries: Example

- $\alpha = 0.05$, $K = 4$ and $\Delta = 0.25$: $C_{WT}(\alpha, K, \Delta) = 2.113$
  - $c_1 = 2.988$
  - $c_2 = 2.513$
  - $c_3 = 2.271$
  - $c_4 = 2.113$

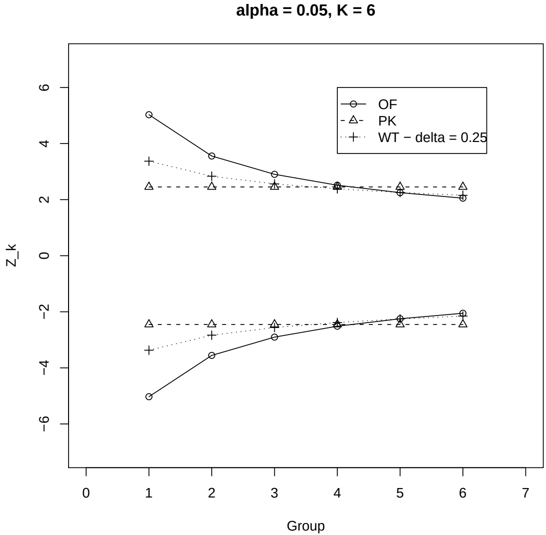
Wang and Tsiatis Boundaries: Example



Wang and Tsiatis Boundaries: Example

- $\alpha = 0.05$, $K = 6$ and $\Delta = 0.25$: $C_{WT}(\alpha, K, \Delta) = 2.154$
  - $c_1 = 3.371$
  - $c_2 = 2.835$
  - $c_3 = 2.562$
  - $c_4 = 2.384$
  - $c_5 = 2.254$
  - $c_6 = 2.154$

Wang and Tsiatis Bounds: Examples



Haybittle-Peto Tests

- Haybittle (1971) and Peto (1976) suggest a simple approach of using a very large critical value for the interim analysis and adjusting hte final analyses to achieve the desired type-I error
- Specifically:
  - $c_k = 3.0$ for all $k = 1, \ldots, K$
  - Find $c_K = C_{HP}(\alpha, K)$ that provides the desired type-I error rate

Haybittle-Peto Test

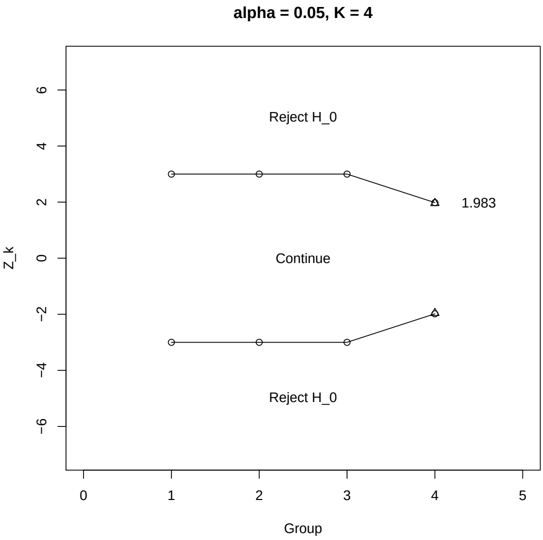
- Formally, the Haybittle-Peto test is as follows:
- For $k = 1, \ldots, K - 1$
  - if $|Z_k| > 3$, stop and reject H_0
  - otherwise, continue to group $k + 1$
- For $k = K$
  - if $|Z_K| > c_{HP}(\alpha, K)$, stop and reject H_0
  - otherwise, stop and fail to reject H_0

Haybittle-Peto Boundaries: Example

Examples

- $\alpha = 0.05$ and $K = 4$: $C_{OF}(\alpha, K) = 1.983$
  - $c_1 = 3$
  - $c_2 = 3$
  - $c_3 = 3$
  - $c_4 = 1.983$

Haybittle-Peto Boundaries: Example

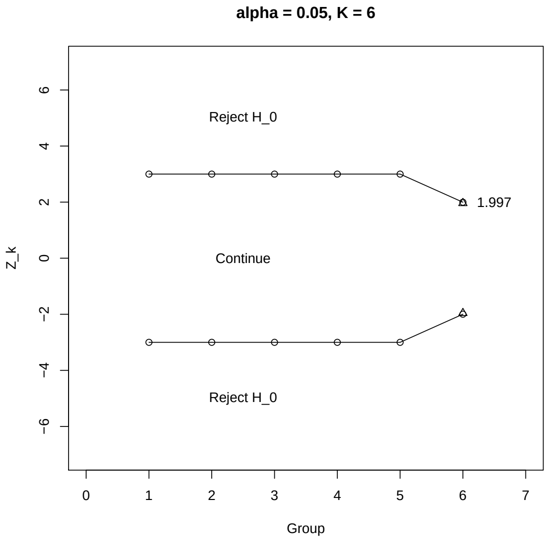


Haybittle-Peto Boundaries: Example

Examples

- $\alpha = 0.05$ and $K = 6$: $C_{OF}(\alpha, K) = 1.997$
  - $c_1 = 3$
  - $c_2 = 3$
  - $c_3 = 3$
  - $c_4 = 3$
  - $c_5 = 3$
  - $c_6 = 1.997$

Haybittle-Peto Boundaries: Example



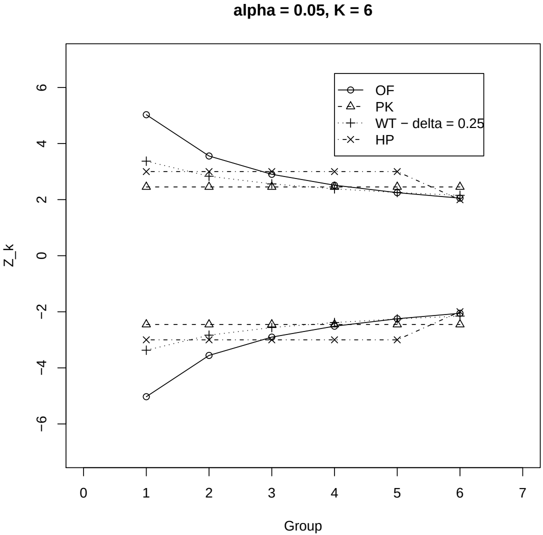
Haybittle-Peto Test: Advantages and Disadvantages

- Advantages
  - Simple!
  - Results in final critical value close to critical value for fixed-sample test
- Disadvantages
  - It is not possible to find C_{HP} that achieves the desired type-I error rate for some combinations of α and K
  - i.e. if α is too small or K is too large

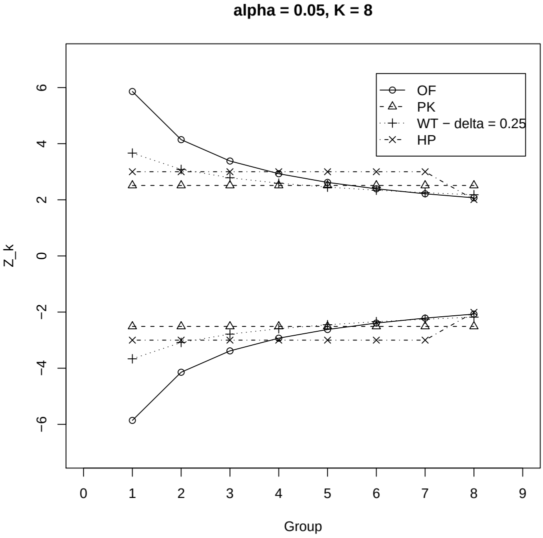
Comparison of Stopping Boundaries

- We have discussed four types of stopping boundaries
  - Pocock boundaries
  - O'Brien-Fleming boundaries
  - Wang and Tsiatis boundaries
  - Haybittle-Peto boundaries

Comparison of Stopping boundaries



Comparison of Stopping boundaries



Comparison of Stopping boundaries

For $\alpha = 0.05$ and $K = 8$

- At $k = 1$, from largest to smallest critical values
  - O'Brien-Fleming
  - Wang and Tsiatis
  - Haybittle-Peto
  - Pocock

Comparison of Stopping boundaries

For $\alpha = 0.05$ and $K = 8$

- At $k = 5$, from largest to smallest critical values
  - Haybittle-Peto
  - O'Brien-Fleming
  - Wang and Tsiatis
  - Pocock

Comparison of Stopping boundaries

For $\alpha = 0.05$ and $K = 8$

- At $k = 8$, from largest to smallest critical values
  - Pocock
  - Wang and Tsiatis
  - O'Brien-Fleming
  - Haybittle-Peto

Comparison of Stopping boundaries

- In general, there is a trade-off between the magnitude of critical values at interim analysis and at study completion
  - Studies with smaller critical values at the interim analyses will have larger critical values at study completion
  - O'Brien-Fleming and Heybittle-Peto boundaries have large critical values at the interim analyses but relatively small critical values at study completion
  - Pocock boundaries have small critical values at the interim analyses but large critical values at study completion

Comparison of Stopping boundaries

- The four stopping boundaries we've considered have the same type-1 error rate
- How do they compare for?
  - Power/maximum sample size
  - Expected sample size

Power and Group Sequential Designs

- Recall that adding interim analyses decreases the power
- We must increase the maximum sample size to achieve the desired power
- The increase in sample size will depend on α, $1-\beta$, K and the type of boundaries being used

Power Example: Pocock Boundaries

- When comparing means for two normally distributed random variables, a sample size of 100 subjets per group provides 80% power to detect a 0.4 standard deviation difference with $\alpha = 0.05$
- Consider a group sequential test with $K = 5$ and Pocock stopping boundaries
- Maximum sample size must be multiplied by 1.229 to achieve 80% power
  - Maximum sample size increased to 123
  - Interim analysis at n = 25, 50, 74 and 99

Maximum Sample Size Comparison

- Maximum Sample Size inflation factor for
  - $\alpha = 0.05$
  - $1 - \beta = 0.80$

Boundary	K = 4	K = 6	K = 8
Pocock	1.202	1.249	1.279
O'Brien-Fleming	1.024	1.032	1.037
Wang and Tsiatis: $\Delta = 0.25$	1.065	1.077	1.084
Haybittle-Peto	1.011	1.019	1.027

Maximum Sample Size Comparison

- Maximum Sample Size inflation factor for
  - $\alpha = 0.05$
  - $1 - \beta = 0.90$

Boundary	K = 4	K = 6	K = 8
Pocock	1.183	1.225	1.252
O'Brien-Fleming	1.022	1.030	1.034
Wang and Tsiatis: $\Delta = 0.25$	1.059	1.071	1.078
Haybittle-Peto	1.010	1.017	1.024

Maximum Sample Size Summary

- Maximum sample size is closely related to final critical value
- Stopping boundaries with smaller final critical values have smaller maximum sample size
- Pocock boundaries require a large maximum sample size
- O'Brien-Fleming and Haybittle-Peto boundaries require only a small increase in sample size
- Wang and Tsiatis boundaries are in the middle depending on Δ

Expected Sample Size

- Sample size is a random variable in a group sequential design
- We can compare designs by the expected sample size
- The expected sample size will depend on α, $1 - \beta$, K, the type of boundaries being used and the true difference between groups

Expected Sample Size Example: Pocock boundaries

- A fixed-sample test requires 100 subjets per group to achieve 80% power to detect a 0.4 standard deviation difference with $\alpha = 0.05$
- Pocock stopping boundaries with $K = 5$ requires a maximum sample size of 123 to achieve 80
- The expected sample size depends on the true difference between groups
  - $\delta = 0$: expected sample size equals 119.8
  - $\delta = 0.2\sigma$: expected sample size equals 110.4
  - $\delta = 0.4\sigma$: expected sample size equals 79.9
  - $\delta = 0.6\sigma$: expected sample size equals 50.1

Expected Sample Size Comparison

- Expected sample size as percentage of sample size required to detect a difference of δ
  - $\alpha = 0.05$
  - $1 - \beta = 0.80$
  - $K = 5$

	True Difference			
Boundary	0	.5 δ	δ	1.5 δ
Pocock	119.8	110.4	79.9	50.1
O'Brien-Fleming	102.1	97.9	81.8	61.9
Wang and Tsiatis: $\Delta = 0.25$	105.8	99.8	78.7	55.2
Haybittle-Peto	101.1	98.8	85.9	61.5

Expected Sample Size Comparison

- Expected sample size as percentage of sample size required to detect a difference of δ
  - $\alpha = 0.05$
  - $1 - \beta = 0.90$
  - $K = 5$

	True Difference			
Boundary	0	.5 δ	δ	1.5 δ
Pocock	117.7	105.2	68.5	41.2
O'Brien-Fleming	101.9	96.1	75.0	54.8
Wang and Tsiatis: $\Delta = 0.25$	105.3	97.0	70.4	47.3
Haybittle-Peto	100.9	97.6	78.8	50.8

Expected Sample Size: Summary

- Pocock boundaries result in dramatic increases in sample size with no or small difference between groups but big savings with large differences
- O'Brien-Fleming and Haybittle-Peto boundaries result in minimal increase in sample size under null but less savings with large differences
- Wang and Tsiatis boundaries are somewhere in between

How to pick K?

- How should we go about picking K when designing a study?
- We know that increasing K will increase the maximum sample size.
- What about the expected sample size?
    - It depends on the stopping boundaries and true difference between groups

Effect of K on Expected Sample Size: Pocock

- Expected sample size as percentage of sample size required to detect a difference of δ
  - $\alpha = 0.05$
  - $1 - \beta = 0.80$

		True Difference			
	Max SS	0	.5 δ	δ	1.5 δ
$K = 2$	111.0	109.4	103.9	85.3	65.0
$K = 3$	116.6	114.3	106.7	81.9	56.0
$K = 4$	120.2	117.5	108.8	80.5	52.2
$K = 5$	122.9	119.8	110.4	79.9	50.1
$K = 10$	130.1	126.3	115.3	79.5	46.2
$K = 15$	133.8	129.7	118.1	80.0	45.1
$K = 20$	136.3	131.9	120.0	80.5	44.7

Effect of K on Expected Sample Size: O'Brien-Fleming

- Expected sample size as percentage of sample size required to detect a difference of δ
    - $\alpha = 0.05$
    - $1 - \beta = 0.80$

		True Difference			
	Max SS	0	.5 δ	δ	1.5 δ
$K = 2$	100.8	100.5	99.0	90.2	71.9
$K = 3$	101.7	101.2	98.3	85.6	68.0
$K = 4$	102.4	101.7	98.0	83.1	64.1
$K = 5$	102.8	102.1	97.9	81.8	61.9
$K = 10$	104.0	103.1	97.9	79.1	58.1
$K = 15$	104.5	103.5	97.9	78.3	56.9
$K = 20$	104.7	103.7	97.9	77.9	56.3

Effect of K on Expected Sample Size: Wang and Tsiatis

- Expected sample size as percentage of sample size required to detect a difference of δ
  - $\alpha = 0.05$
  - $1 - \beta = 0.80$
  - $\Delta = 0.25$

		True Difference			
	Max SS	0	.5 δ	δ	1.5 δ
$K = 2$	103.8	103.0	99.7	86.0	66.1
$K = 3$	105.4	104.4	99.6	82.0	66.1
$K = 4$	106.5	105.2	99.7	79.9	57.1
$K = 5$	107.2	105.8	99.8	78.7	55.2
$K = 10$	108.9	107.3	100.2	76.2	51.3
$K = 15$	109.7	108.0	100.4	75.4	50.0
$K = 20$	110.1	108.3	100.5	75.0	49.4

Effect of K on Expected Sample Size: Haybittle-Peto

- Expected sample size as percentage of sample size required to detect a difference of δ
  - $\alpha = 0.05$
  - $1 - \beta = 0.80$

		True Difference			
	Max SS	0	.5 δ	δ	1.5 δ
$K = 2$	100.3	100.2	99.2	92.5	75.7
$K = 3$	100.7	100.5	98.9	89.2	68.1
$K = 4$	101.1	100.8	98.8	87.2	64.0
$K = 5$	101.5	101.1	98.8	85.9	61.5
$K = 10$	103.3	102.5	99.3	83.0	55.6
$K = 15$	104.8	103.7	100.0	81.8	53.3
$K = 20$	106.1	104.8	100.7	81.3	51.9

Effect of K on Expected Sample Size: Summary

- In general, there is little benefit of more than 5 stopping times
- The decrease in expected sample size is minimal after $K = 5$ and the expected sample size actually increases in some cases
- The maximum sample size increases dramatically for the Pocock boundaries
- Excessively large number of stopping times becomes unwieldy

Software

- The best (free) software I've found for calculating group sequential stopping boundaries is the gsDesign package in R
- The gsDesign package will calculate Pocock, O'Brien-Fleming and Wang and Tsiatis stopping boundaries (Haybittle-Peto boundaries are not supported)
- The gsDesign package will calculate stopping boundaries, maximum sample size and expected sample size under the null and alternative hypothesis

Calculating Stopping Boundaries using the gsDesign package

The primary function for calculating group sequential stopping boundaries is the gsDesign function. Key inputs include:

- k: the number of stopping times
- test.type: one-sided, two-sided symmetric, two-sided asymmetric, etc. For now, we're only using two-sided symmetric, which is coded as 2
- alpha: one-sided type-I error rate (i.e. use 0.025 for a two-sided alpha of 0.05)
- beta: 1 - power

Calculating Stopping Boundaries using the gsDesign package

- n.fix: sample size for a fixed-sample design
- sfu: boundary type, options include ”Pocock”, ”OF” for O’Brien-Fleming and ”WT” for wang and tsiatis. Other options are available and will be explained later in the semester
- sfupar: boundary parameter. This is Δ for the Wang and Tsiatis boundaries and is not used for Pocock or O’Brien-Fleming boundaries

Calculating Stopping Boundaries: Example

Consider the following clinical trial to evaluate a new blood pressure medication to the standard of care

- We will consider a randomized clinical trial with subjects randomized in a 1:1 ratio to the new treatment and standard of care
- The final outcome will by systolic blood pressure at 12 months
- We expect a difference of 5 mmHg
- The standard deviation is known to be 15 mmHg
- With $\alpha = 0.05$ and 90% power, the sample size for a fixed-sample design is 190 subjects per group

Calculating Stopping Boundaries: Example

Boundary type	K	Max SS	$E(SS \mid H_0)$	$E(SS \mid H_A)$
Fixed-sample	1	190	190	190
Pocock	5	230	223.6	130.1
O'Brien-Fleming	5	196	193.6	142.5
Wang and Tsiatis ($\Delta = 0.25$)	5	203	200.0	133.7

Calculating Stopping Boundaries: Example

- The Pocock boundaries have the potential of a much larger sample size than the other two options
- O'Brien-Fleming require 30 less subjects under the null but only 12.5 more under the alternative
- Wang and Tsiatis is even better:
  - 23.6 less under the null
  - 3.6 more under the alternative

Sequential Monitoring Example: MRFIT Study

- Randomized control trial on the prevention coronary heart disease (CHD) mortality
- Men aged 35 - 57 years old
- No definitive evidence of CHD at baseline

MRFIT treatment

- Men were randomized to either usual care or experimental intervention
  - intervention: dietary counseling, smoking cessation counseling and stepped-care hypertension medication
  - usual care: treatment by usual primary care physician

MRFIT outcomes

- Men returned every year for six years and the following outcomes were recorded
  - medication use
  - systolic blood pressure
  - diastolic blood pressure
  - triglycerides
  - HDL cholesterol
  - LDL cholesterol

Sequential monitoring in MRFIT

- We will focus on a subset of 800 (400 per group) and the outcomes systolic and diastolic blood pressure at 24 months
- Consider group sequential designs with 5 stopping times at 80, 160, 240, 320 and 400 subjects per group
- We have two outcomes and will use a Bonferroni adjusted type-1 error rate of 0.025 for each outcome
- We will consider Pocock, O'Brien-Fleming and Wang and Tsiatis boundaries

Stopping Boundaries

K	n	PK	OF	WT ($\Delta = 0.25$)
1	160	2.67	5.15	3.57
2	320	2.67	3.64	3.01
3	480	2.67	2.97	2.72
4	640	2.67	2.58	2.53
5	800	2.67	2.30	2.39

Test statistic

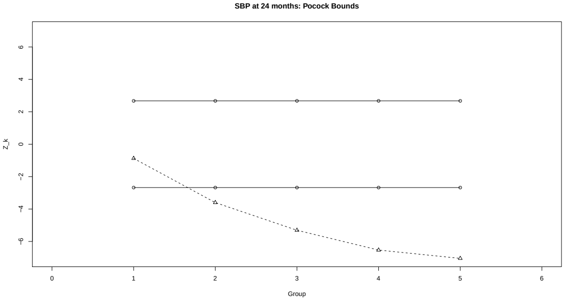
- For our example, we will assume that the standard deviation for SBP is 14 mmHg and standard deviation of DBP is 8.5, both known
- Out test statistic, Z_k will be defined as before

$$Z_k = \frac{\bar{X}_{trt} - \bar{X}_{ctl}}{\sqrt{2\sigma^2 / n_k}}$$

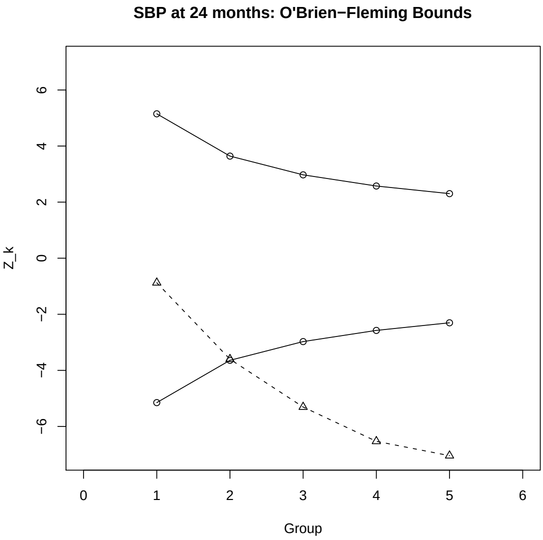
Sequential test statistic: SBP

K	n per group	$\bar{X}_{trt}$	$\bar{X}_{ctl}$	Z_k
1	80	123.7	125.7	-0.88
2	160	122.0	127.6	-3.60
3	240	121.4	128.2	-5.31
4	320	121.8	129.0	-6.53
5	400	121.8	128.8	-7.04

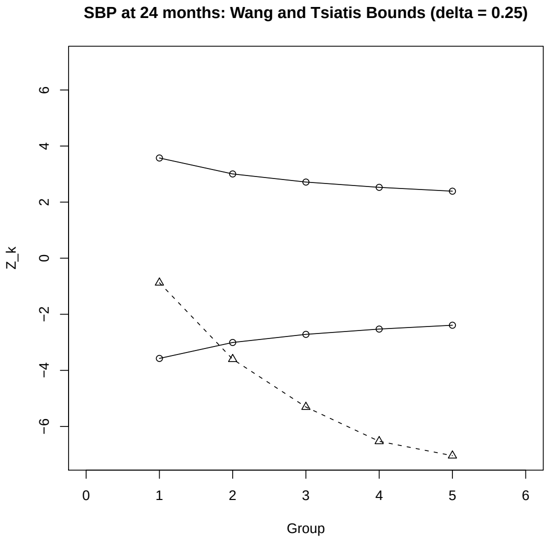
Sequential Monitoring of SBP: Pocock



Sequential Monitoring of SBP: O'Brien-Fleming



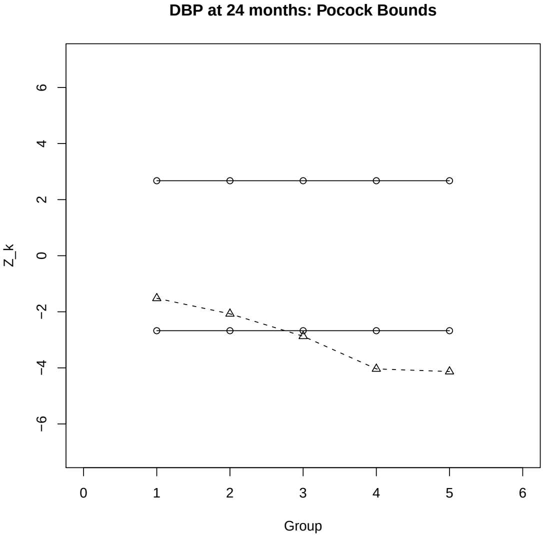
Sequential Monitoring of SBP: Wang and Tsiatis



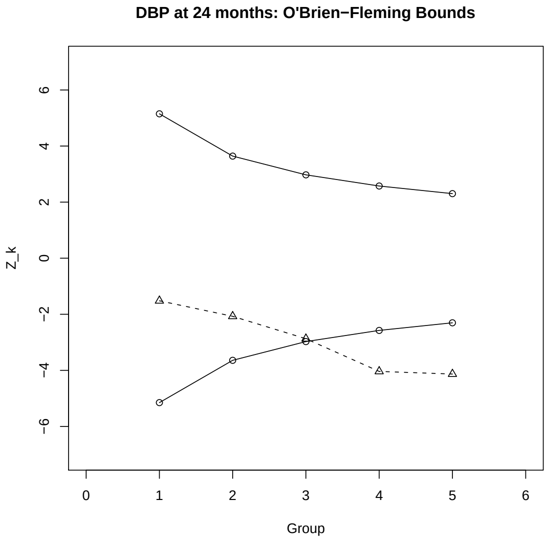
Sequential test statistic: DBP

K	n per group	$\bar{X}_{trt}$	$\bar{X}_{ctl}$	Z_k
1	80	84.3	87.6	-1.52
2	160	83.6	86.9	-2.07
3	240	82.7	86.4	-2.87
4	320	82.5	87.0	-4.04
5	400	82.6	86.6	-4.13

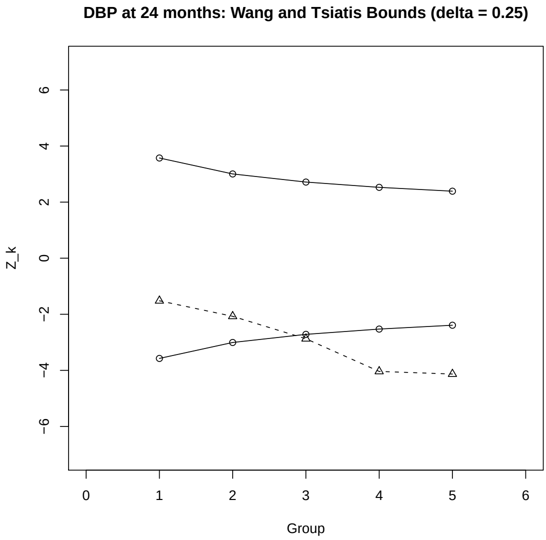
Sequential Monitoring of DBP: Pocock



Sequential Monitoring of DBP: O'Brien-Fleming



Sequential Monitoring of DBP: Wang and Tsiatis



Sequential Monitoring of MRFIT: Conclusions

- We reject the null hypothesis for both endpoints
- Reject null for SBP after second stopping time for Pocock and Wang and Tsiatis boundaries and after the third for O'Brien-Fleming boundaries
- Reject null for DBP after third stopping time for Pocock and Wang and Tsiatis boundaries and after the fourth for O'Brien-Fleming boundaries
- Assuming that we would stop the trial only if both are rejected, we would stop after the third interim analysis for Pocock and Wang and Tsiatis boundaries and after the fourth for O'Brien-Fleming
- Therefore, we would require a sample size of 480 (240/group) for the Pocock and Wang and Tsiatis boundaries and 640 (320/group) for the O'Brien-Fleming boundaries

Brownian Motion

Let $W(t)$ be a continuous Gaussian process defined on $0 \leq t \leq 1$.
$W(t)$ is a Brownian Motion if:

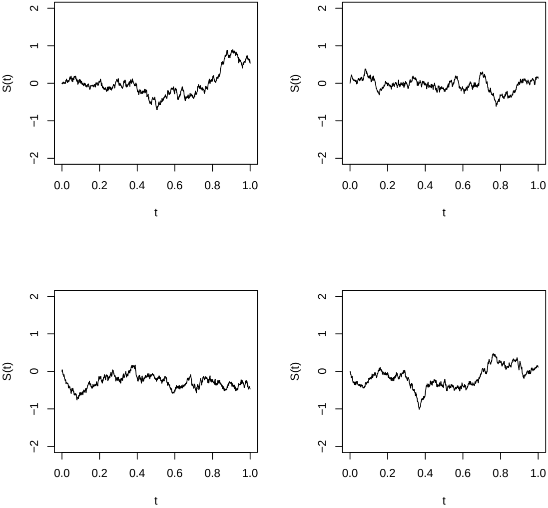
- $W(0) = 0$
- $E\left[W\left(t\right)\right] = t$ for all $t \in [0, 1]$
- $Cov\left[W\left(t_1\right), W\left(t_2\right)\right] = t_1$ for $0 \leq t_1 \leq t_2 \leq 1$

Properties of Brownian Motion

A nice property of Brownian motion is that Brownian Motion has independent increments. That is,

- $W(t_2) - W(t_1)$ and $W(t_4) - W(t_3)$ are independent for $0 \le t_1 \le t_2 \le t_3 \le t_4 \le 1$

Samples Paths for Brownian Motion



Partial Sum Process

Let $X_1, X_2, \ldots, X_n$ be i.i.d. random variables with mean 0 and variance 1. Define the partial sum process, $S_n(t)$ by

$$S_n(t) = \frac{1}{\sqrt{n}} \sum_{i=1}^{[nt]} X_i$$

where $[nt] = k$ such that,

$$\frac{k}{n} \leq t < \frac{k+1}{n}$$

Partial Sum Process and Brownian Motion

You can show that

$$S_n(t) \rightarrow_d W(t)$$

This a helpful result that allows us to derive the joint distribution of Z_k.