

Part II

Second-generation *p*-values: statistical properties, false discovery rates, equivalence tests, and variable selection

Jeffrey D. Blume, and Megan H. Murray,

PhD

PhD

School of Data Science

Statistician at

University of Virginia

Eli Lilly and Company

Outline

- Statistical Properties
- False Discovery Rates
- Study Planning
- Comparison to Equivalence Tests
- SGPV Variable Selection (if we have time)

Connecting SGPV and Properties

- Three Inferential Categories
 1. $p_\delta = 0 \Rightarrow$ data **incompatible** with null
 2. $p_\delta = 1 \Rightarrow$ data **compatible** with null
 3. $0 < p_\delta < 1 \Rightarrow$ data are **inconclusive**
- Three ‘error’ rates
 1. $P(p_\delta = 0 | H_0)$ when H is null
 2. $P(p_\delta = 1 | H_1)$ when H is not null
 3. $P(0 < p_\delta < 1 | H)$ when H is either
- Assume H makes statements about a parameter θ
- Large sample setting

ASA SGPV Short Course

Blume and Murray, 2022

Second-generation p -value

- Statistical properties in TAS & PLOS One
- Retains strict error control

Evidential Metric	What it measures	SGPV
1	Summary measure	SGPV (p_δ)
2	Operating characteristics	$P(p_\delta = 0 H_0)$ $P(p_\delta = 1 H_1)$ $P(0 < p_\delta < 1 H)$
3	False discovery rates	$P(H_0 p_\delta = 0)$ $P(H_1 p_\delta = 1)$

ASA SGPV Short Course

Blume and Murray, 2022

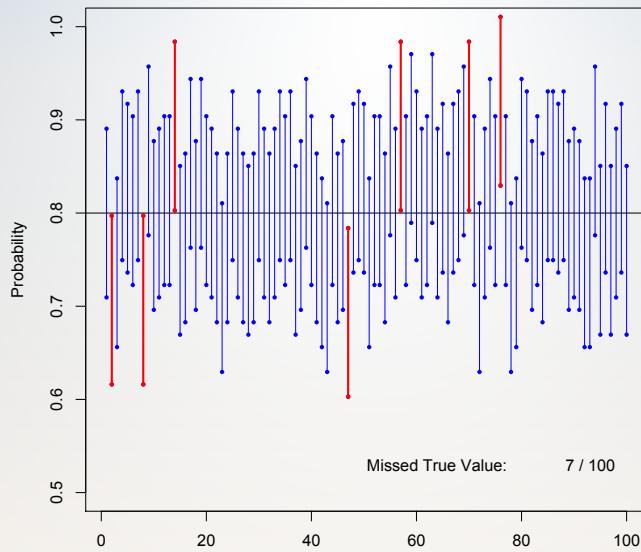
The Intuition

- SGPV error rates converge to zero as the sample size grows.
- Why? The indifference zone is responsible for this.
- Example: 95% Confidence Intervals
- Increasing the sample size...
 - Does not change the miss rate (!)
 - Reduces the width of the CI
 - Reduces the amount by which the CIs miss the truth (!)
- The complement of the coverage probability bounds the SGPV error rates

ASA SGPV Short Course

Blume and Murray, 2022

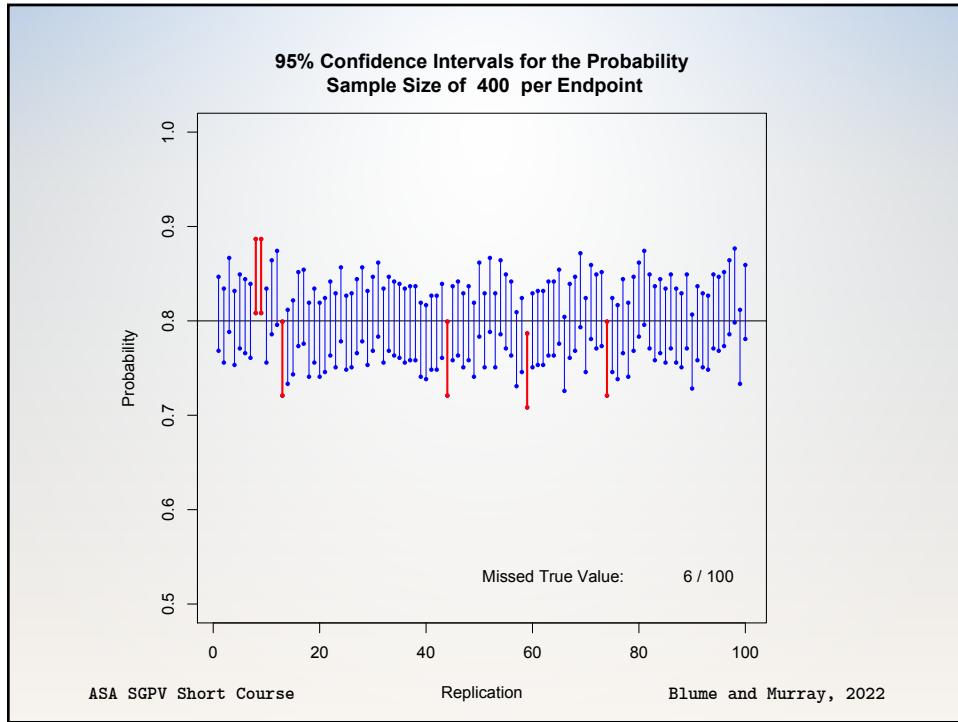
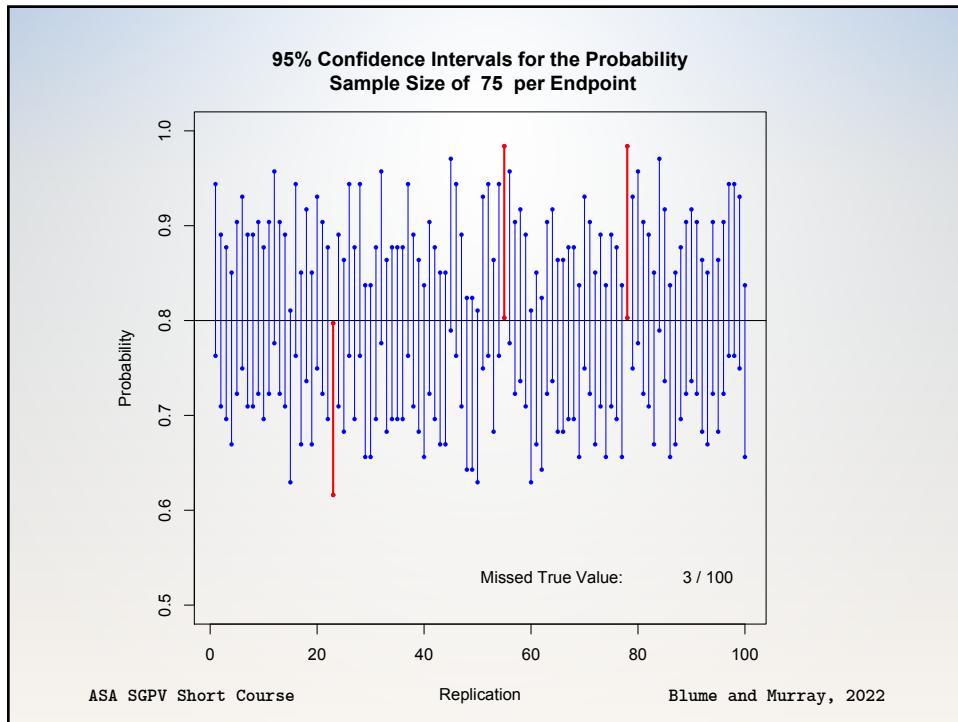
95% Confidence Intervals for the Probability
Sample Size of 75 per Endpoint

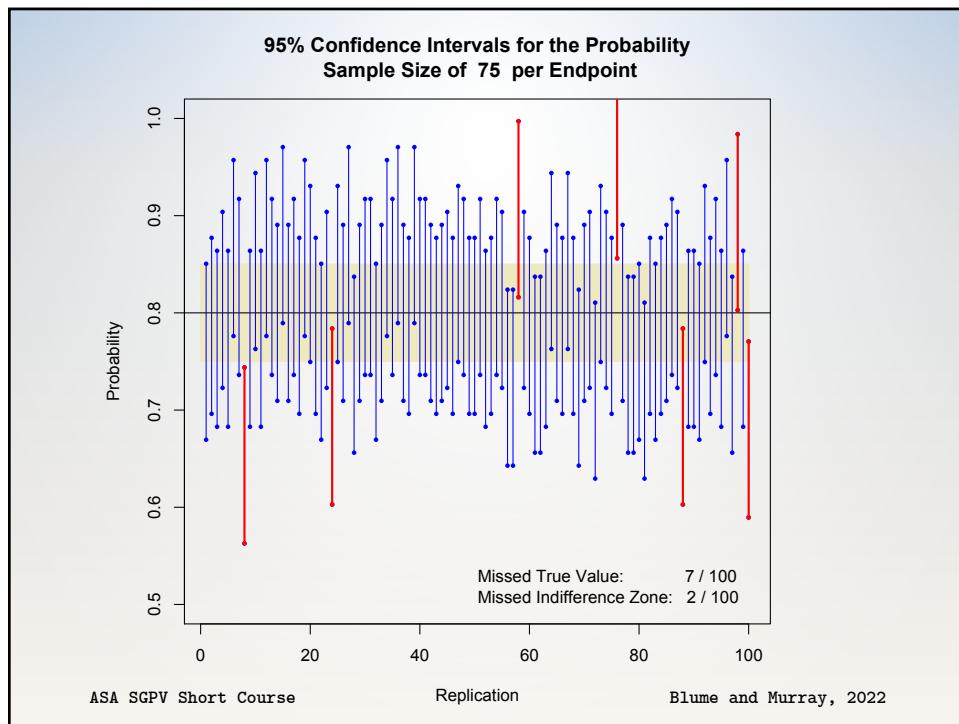
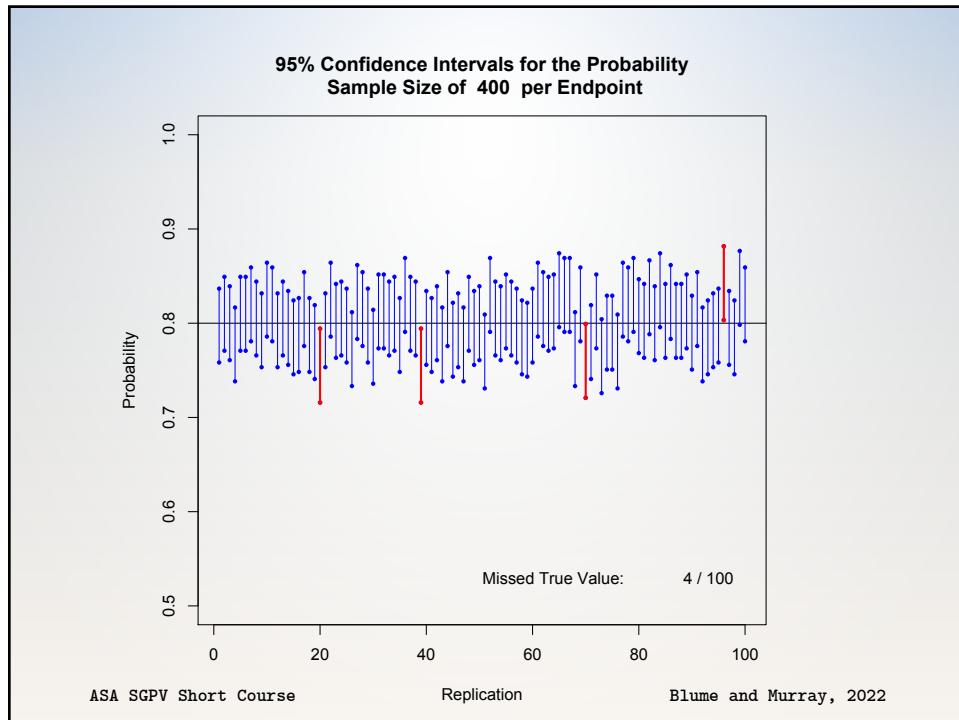


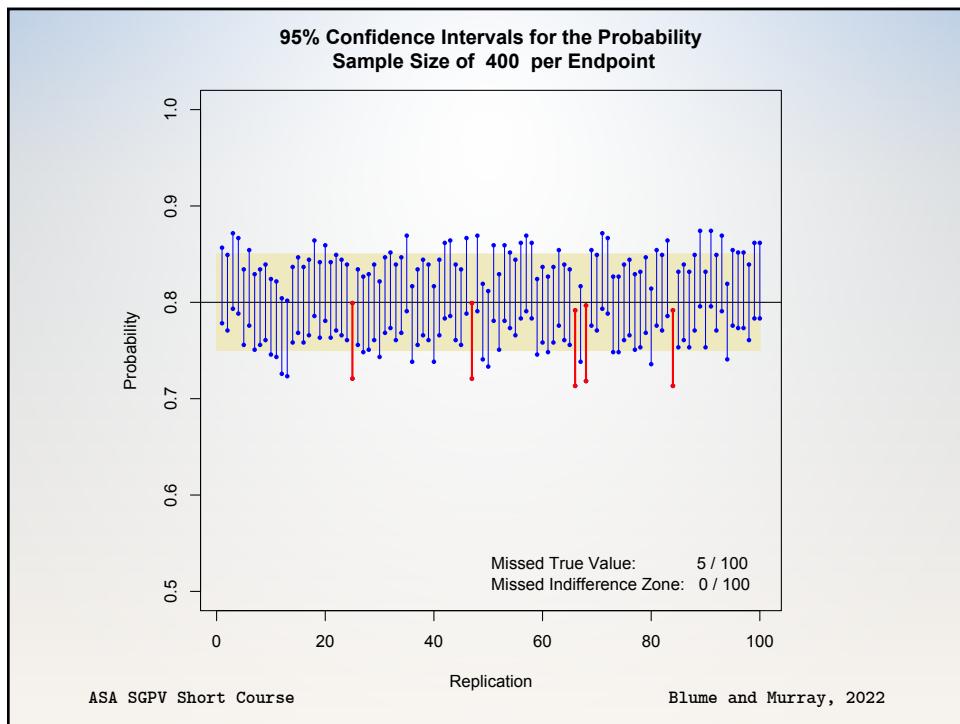
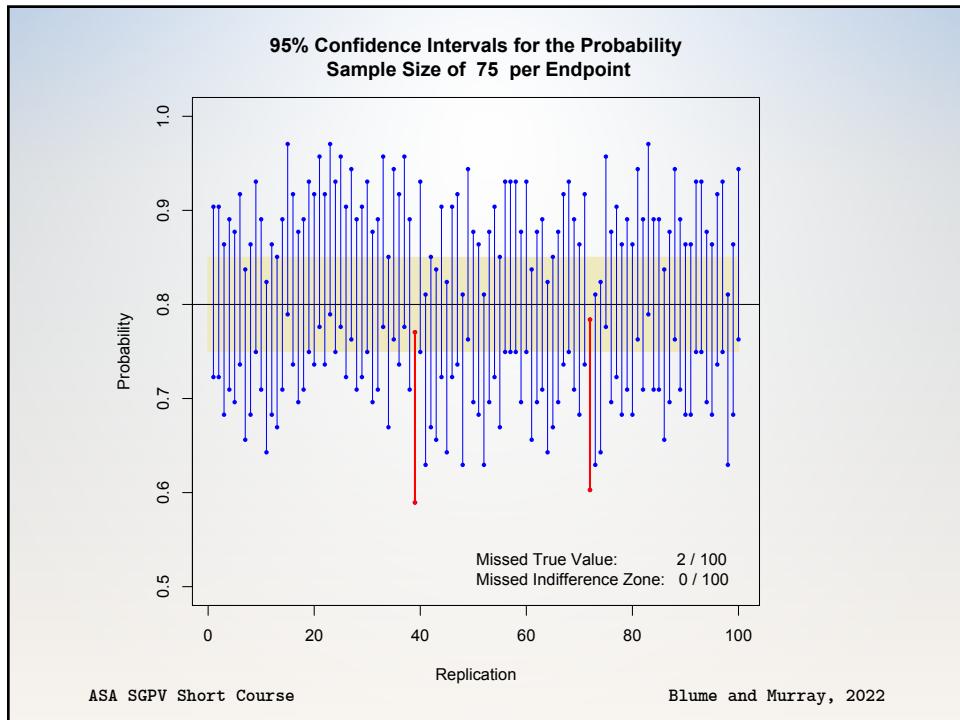
ASA SGPV Short Course

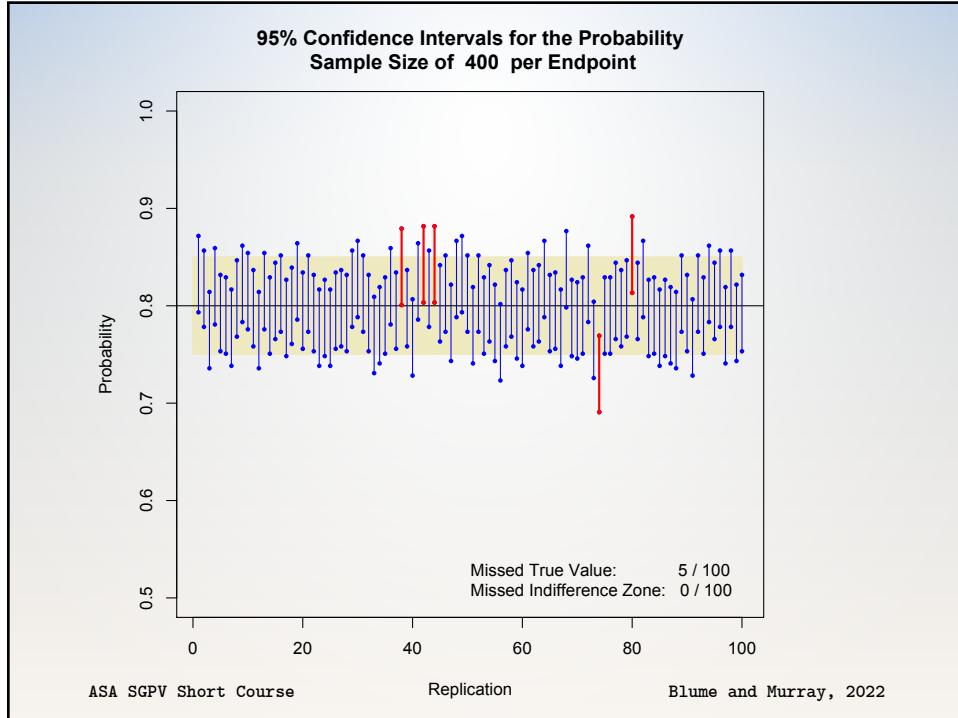
Replication

Blume and Murray, 2022









Statistical Properties

Suppose interval I has coverage probability $1-\alpha$, then

Three ‘Error’ Rates

1. $P(p_\delta = 0 | H_0) \leq \alpha$ and $\rightarrow 0$ as $n \rightarrow \infty$
2. $P(p_\delta = 1 | H_1) \leq \alpha$ and $\rightarrow 0$ as $n \rightarrow \infty$
3. $P(0 < p_\delta < 1 | H)$ controlled through sample size

Will examine
these first

Two False Discovery Rates

1. $P(H_0 | p_\delta = 0)$
2. $P(H_1 | p_\delta = 1)$

Will graph to
illustrate

Statistical Properties

- How often are the data incompatible with null?
- Examine $P(p_\delta = 0|\theta)$ as θ varies
 - Power function
- This probability
 - converges to one for alternatives not near the edge of interval null
 - converges to zero for null hypotheses not near the edge of the null set
 - converges to alpha for hypotheses approaching or on the edge of the null set

ASA SGPV Short Course

Blume and Murray, 2022

‘Power’ Function

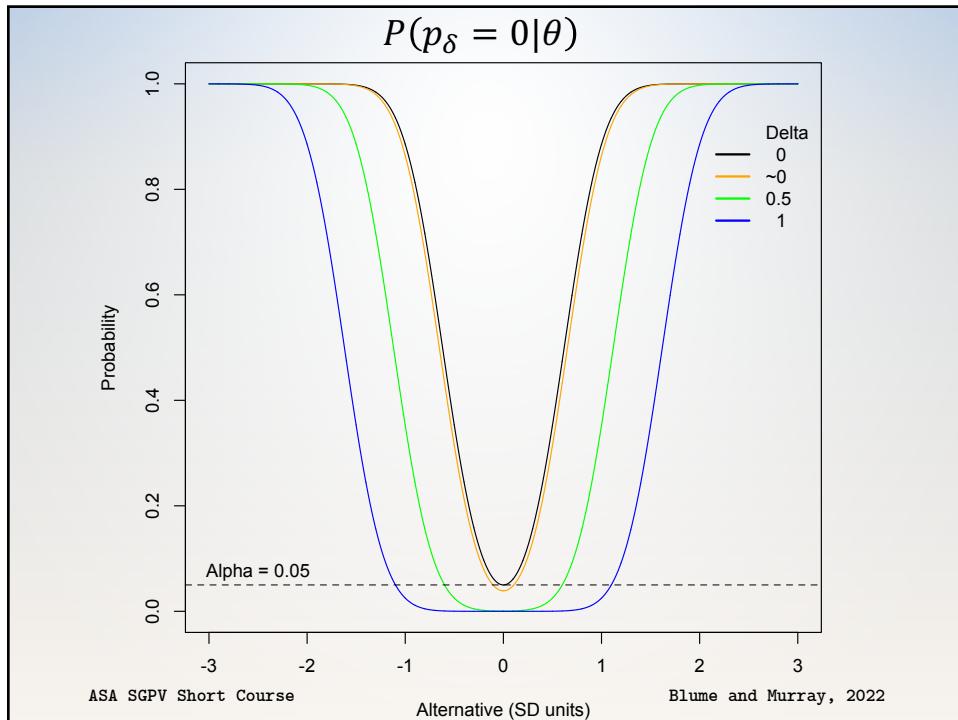
- θ_0 : point null, σ : standard deviation
- δ : half-width of indifference zone

$$\begin{aligned} P(p_\delta = 0|\theta) &= \Phi \left[\frac{\sqrt{n}(\theta_0 - \theta)}{\sigma} - \frac{\sqrt{n}\delta}{\sigma} - Z_{\alpha/2} \right] \\ &\quad + \Phi \left[-\frac{\sqrt{n}(\theta_0 - \theta)}{\sigma} - \frac{\sqrt{n}\delta}{\sigma} - Z_{\alpha/2} \right] \end{aligned}$$

$$P_{\theta_0}(p_\delta = 0|\theta_0) = 2\Phi \left[-\frac{\sqrt{n}\delta}{\sigma} - Z_{\alpha/2} \right]$$

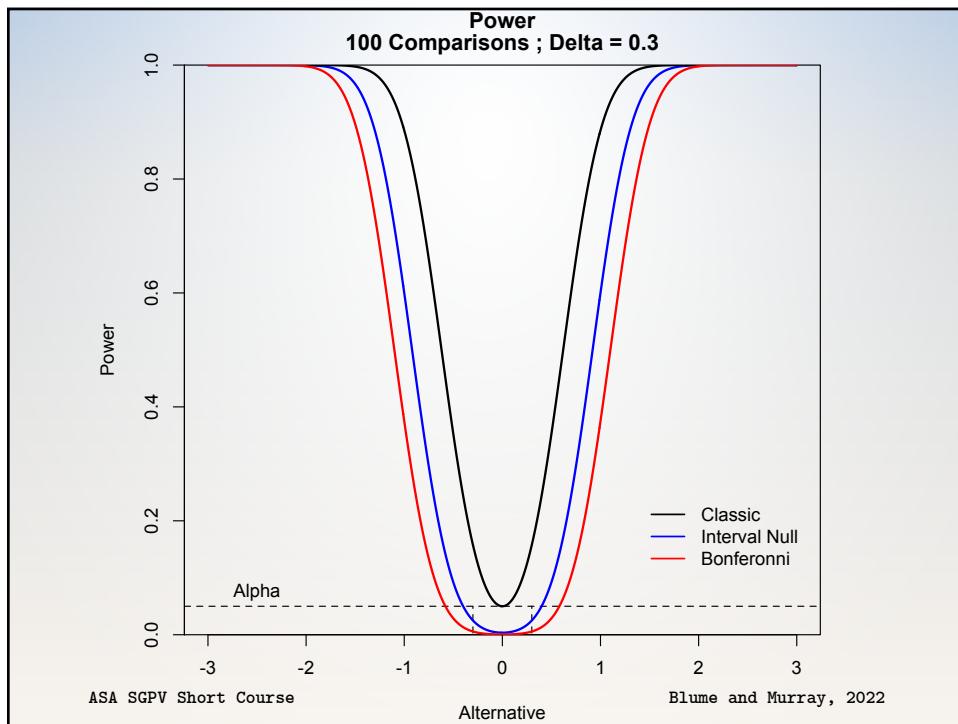
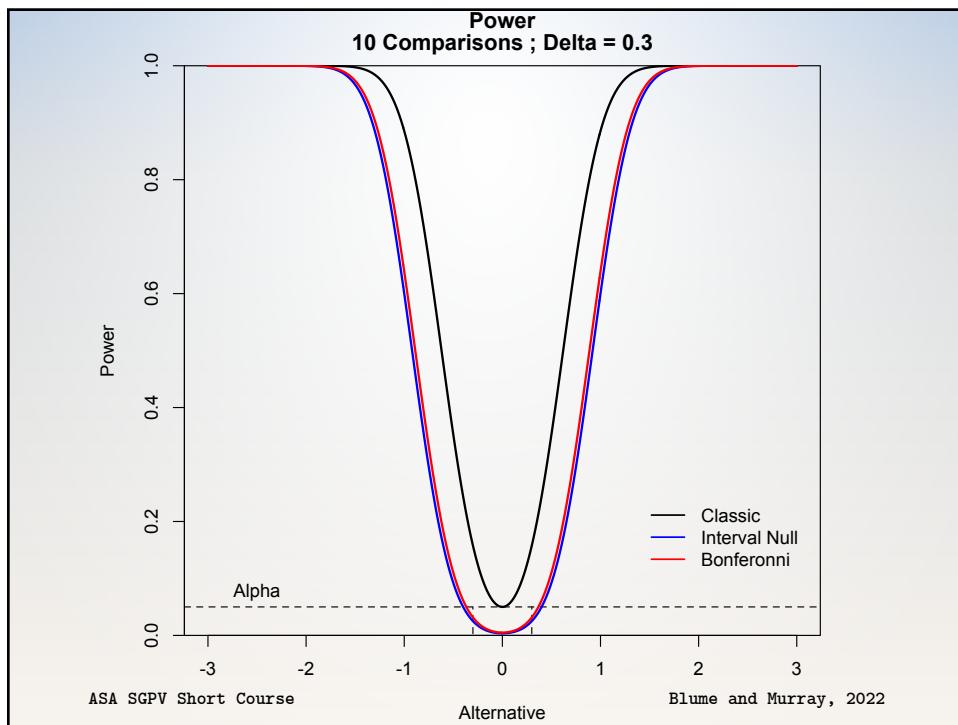
ASA SGPV Short Course

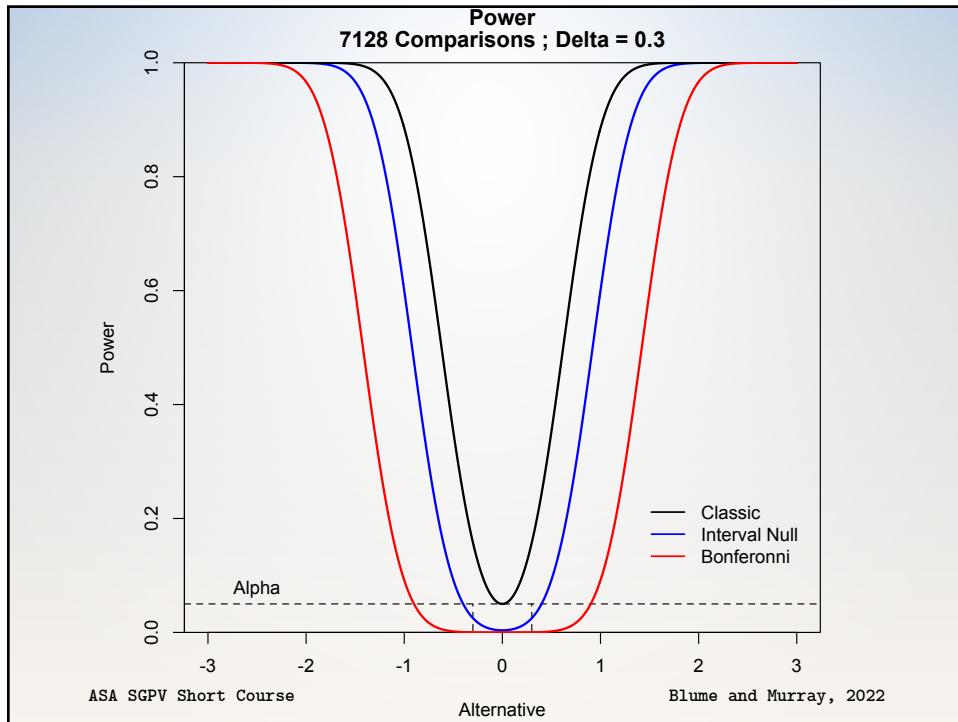
Blume and Murray, 2022



Compare with standard methods

- Second-generation p -value vs. Bonferroni correction
 - Adjusted for $\{10, 100, 7128\}$ comparisons
 - Leukemia data example
- Remember SGPV are not adjusted for comparisons
- Discuss?





Compatible with Null

- How often are the data compatible with null?
- Examine $P(p_\delta = 1|\theta)$ when θ is null or practically null
 - Essentially opposite of power function
- Sample size must be large enough to allow the null interval to contain the interval estimate
- This probability
 - converges to zero or one quickly for alternatives not near the edge of interval null

'Null Power' Function

- How often are the data compatible with null?
- Sample size must be large enough to allow null interval to contain the interval estimate, so $(\delta > Z_{\alpha/2}/\sqrt{n})$ or $(\sqrt{n} > Z_{\alpha/2}/\delta)$
- This probability converges to 0 or 1 quickly

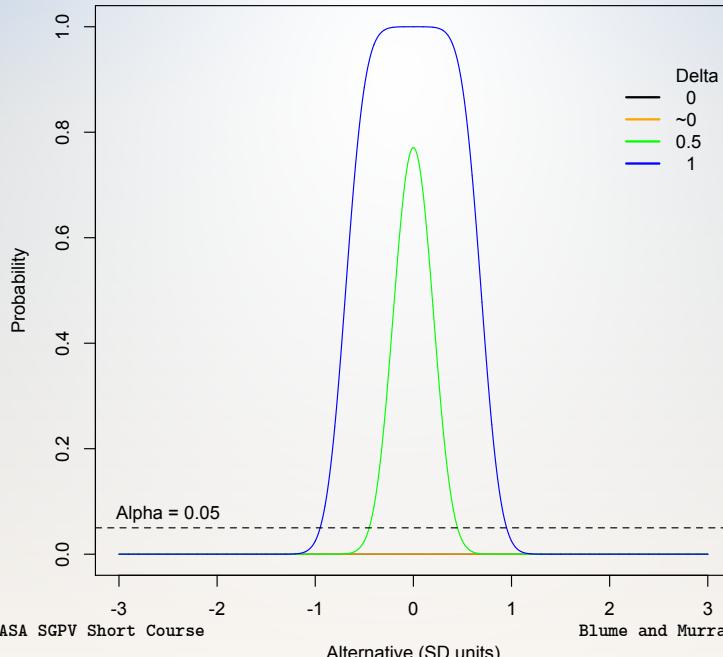
$$P(p_\delta = 1|\theta) = \Phi \left[\frac{\sqrt{n}(\theta_0 + \delta)}{\sigma} - \frac{\sqrt{n}\theta}{\sigma} - Z_{\alpha/2} \right] - \Phi \left[\frac{\sqrt{n}(\theta_0 - \delta)}{\sigma} - \frac{\sqrt{n}\theta}{\sigma} + Z_{\alpha/2} \right]$$

when $\delta > Z_{\alpha/2}/\sqrt{n}$

ASA SGPV Short Course

Blume and Murray, 2022

$$P(p_\delta = 1|\theta)$$



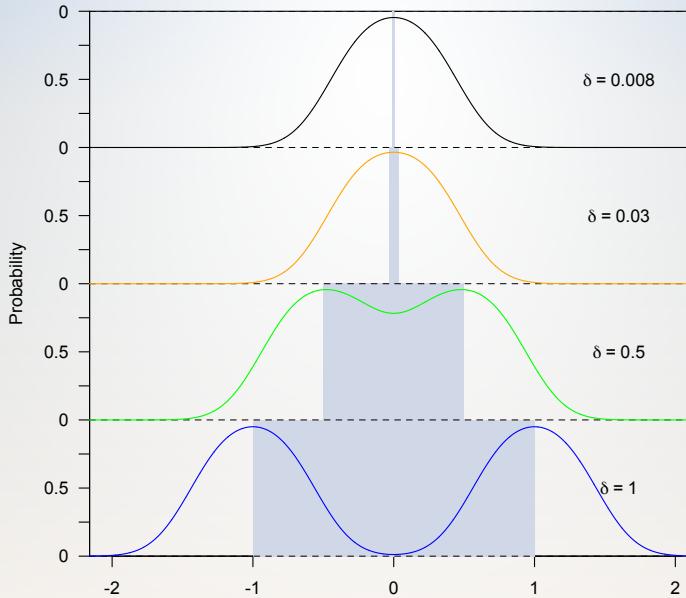
Probability of Inconclusive Data

- How often are the data inconclusive?
- Examine $P(0 < p_\delta < 1|\theta)$ for various θ
- This probability
 - drives sample size projections
 - is maximized when H is near the interval null edge
 - decreases quickly as H moves away from edge of null
- $P(0 < p_\delta < 1|\theta) = 1 - P(p_\delta = 0|\theta) - P(p_\delta = 1|\theta)$

ASA SGPV Short Course

Blume and Murray, 2022

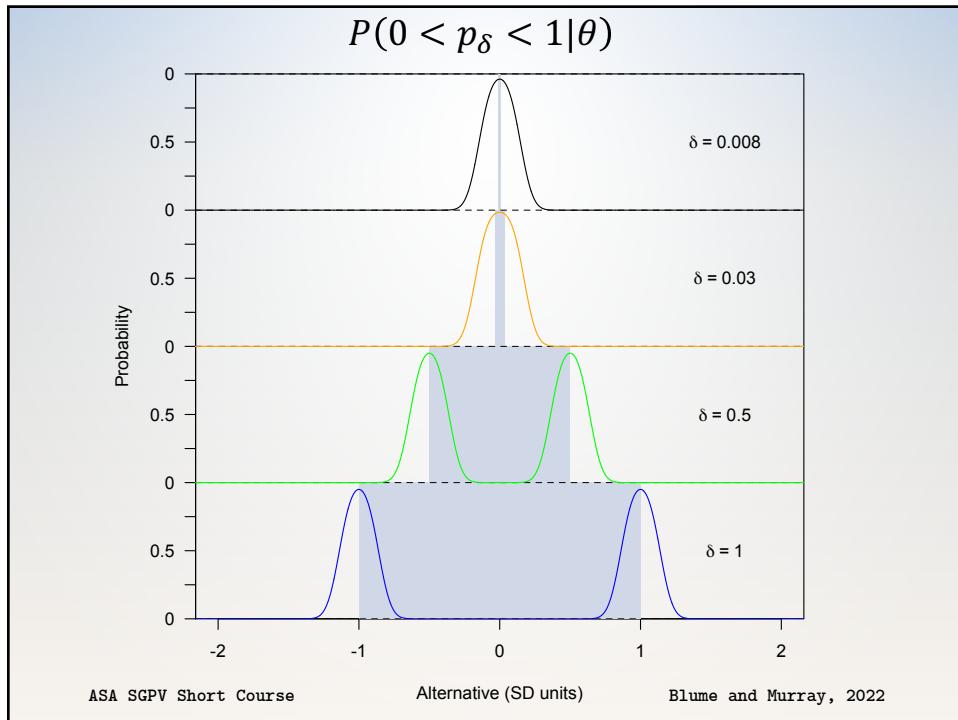
$$P(0 < p_\delta < 1|\theta)$$



ASA SGPV Short Course

Alternative (SD units)

Blume and Murray, 2022



Remarks

- Second-generation p -values...
 - Has three ‘Error’ rates
 - Allows Type I and II rate to converge to zero
 - Control changes of inconclusive results
 - Controls error rate using *science*
 - Reduces the false discovery rate (next section)

- Anchoring the scale of the effect size...
 - Eliminates most Type I Errors
 - Improves scientific translation of statistical model

Statistical Properties

Suppose interval I has coverage probability $1-\alpha$, then

Three ‘Error’ Rates

1. $P(p_\delta = 0 | H_0) \leq \alpha$ and $\rightarrow 0$ as $n \rightarrow \infty$
2. $P(p_\delta = 1 | H_1) \leq \alpha$ and $\rightarrow 0$ as $n \rightarrow \infty$
3. $P(0 < p_\delta < 1 | H)$ controlled through sample size

Two False Discovery Rates

1. $P(H_0 | p_\delta = 0)$
2. $P(H_1 | p_\delta = 1)$

ASA SGPV Short Course

Blume and Murray, 2022

False Discovery Rates

- FDR for 5 SGPV=0 findings; computed under various null and alternative configurations (w/ flat prior).

SNP ID	SGPV rank	p-value rank	OR	1/8 SI lower limit	1/8 SI upper limit	FDR ₁	FDR ₂	FDR ₃
kgp4568244_C	1	133	0.10	0.03	0.37	2.9%	17.1%	3.3%
kgp8051290_G	13	2002	15.58	1.95	124.68	4.3%	30.3%	4.9%
kgp4497498_A	28	255	4.37	1.80	10.64	2.5%	8.6%	3.1%
rs3123636_G	423	1	1.39	1.26	1.55	0.01%	0.1%	0.4%
kgp7460928_G	1443	3310	1.78	1.11	2.87	2.4%	2.0%	3.0%

ASA SGPV Short Course

Blume and Murray, 2022

False discovery rates

- Impact of $\alpha=0.05$ vs $\alpha=0.05/7128$ (7128 comparisons)

- False Discovery Rate (**FDR**)

$$P(H_0|p < \alpha) = \left[1 + \frac{(1 - \beta)}{\alpha} r \right]^{-1}$$

- False Confirmation Rate (**FCR**)

$$P(H_1|p > \alpha) = \left[1 + \frac{(1 - \alpha)}{\beta} \frac{1}{r} \right]^{-1}$$

$$r = P(H_1)/P(H_0)$$

ASA SGPV Short Course

Blume and Murray, 2022

False discovery rates

- Second-generation p -values

- False Discovery Rate (**FDR**)

$$P(H_0|p_\delta = 0) = \left[1 + \frac{P(p_\delta = 0|H_1)}{P(p_\delta = 0|H_0)} r \right]^{-1}$$

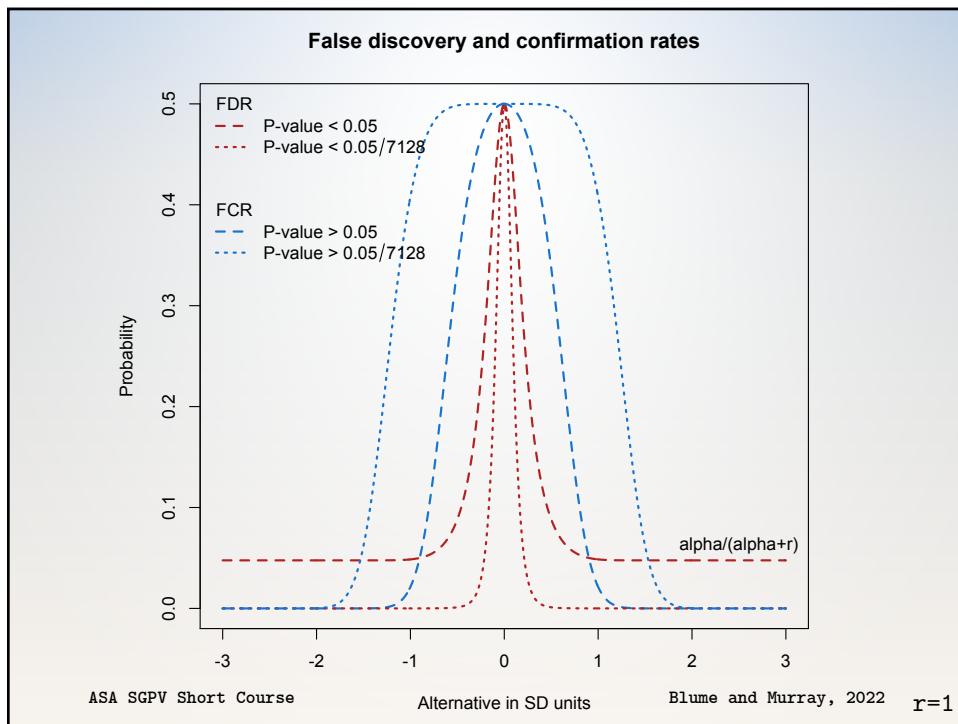
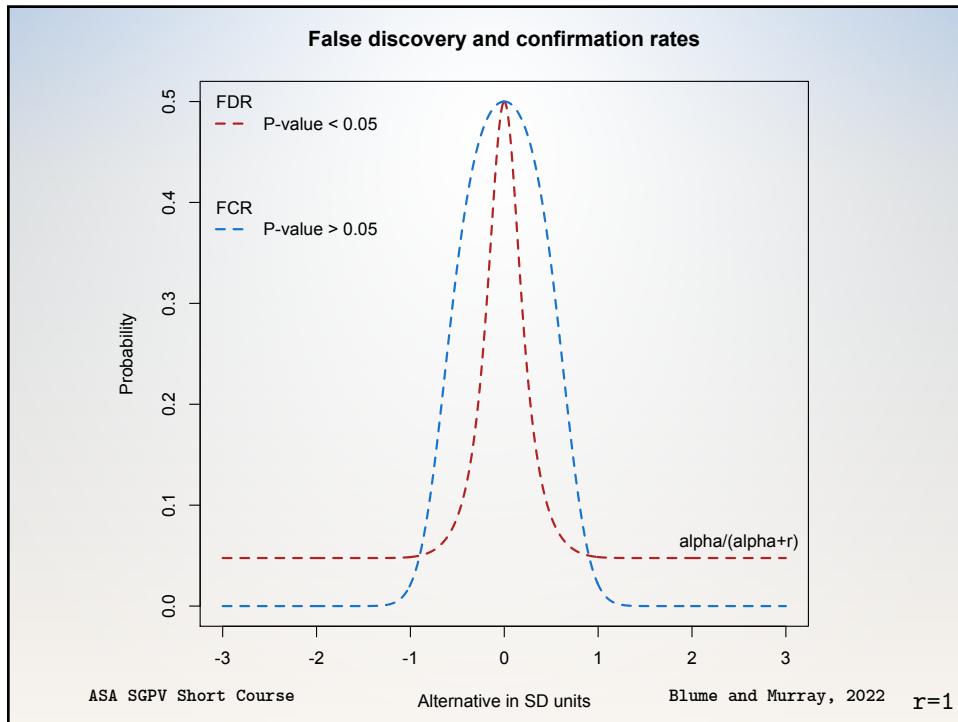
- False Confirmation Rate (**FCR**)

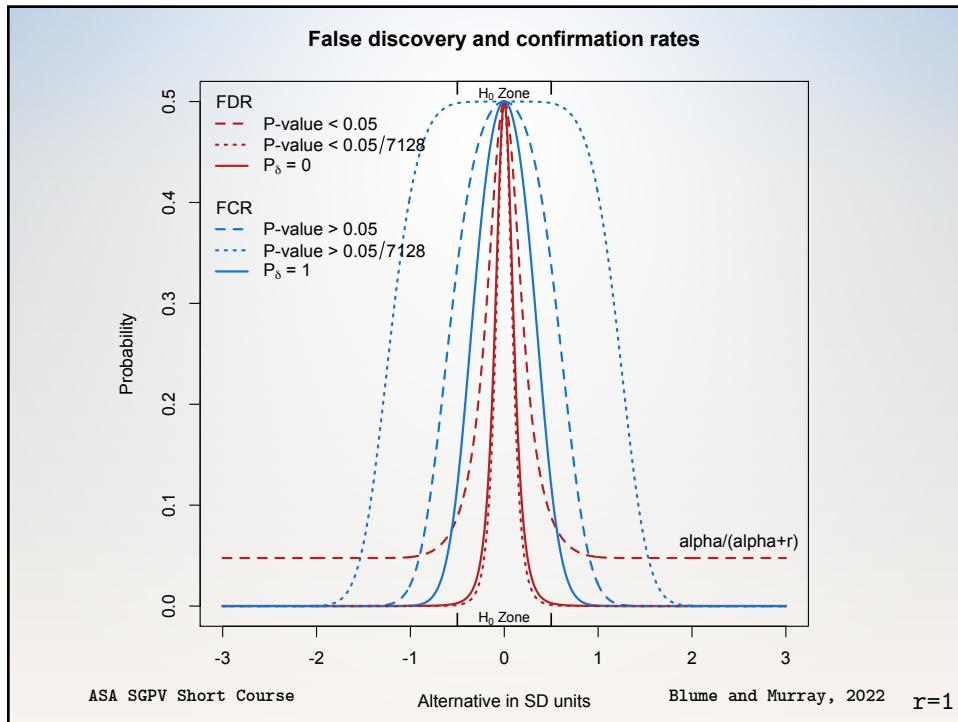
$$P(H_1|p_\delta = 1) = \left[1 + \frac{P(p_\delta = 1|H_0)}{P(p_\delta = 1|H_1)} \frac{1}{r} \right]^{-1}$$

$$r = P(H_1)/P(H_0)$$

ASA SGPV Short Course

Blume and Murray, 2022





FDR R Packages

- SGPVs
 - Valerie Welty
 - `sgpv::fdrisk()`
 - This function computes the false discovery risk (sometimes called the "empirical bayes FDR") for a second-generation p -value of 0, or the false confirmation risk for a second-generation p -value of 1.
- Raw p -values
 - `FDRestimation::p.fdr()`
 - This function computes FDRs and Method Adjusted p -values.
 - Methods include: Benjamini-Hochberg, Benjamini-Yekutieli, Bonferroni, Holm, Hochberg, and Sidak.

Study Planning for SGPVs

- Purpose: Evaluate how different techniques of setting the indifference zone effect the SGPV study properties.
- What options does a collaborator have when they are uncertain of the indifference zone?

ASA SGPV Short Course

Blume and Murray, 2022

Traditional p-value vs. SGPV

		Null is true		Alternative is true	
		Traditional p-value		Traditional p-value	
Inference	Data are consistent with the alternative	Type I error $P_{\theta_0} (p \leq 0.05)$		Power or (1 - Type II error) $P_{\theta_1} (p \leq 0.05)$	
	Data are inconclusive	1 - Type I error $P_{\theta_0} (p > 0.05)$		Type II error $P_{\theta_1} (p > 0.05)$	
	Data are consistent with the null	NA		NA	

ASA SGPV Short Course

Blume and Murray, 2022

Traditional p-value vs. SGPV

		Null is true		Alternative is true	
		SGPV		SGPV	
Inference	Data are consistent with the alternative		β_0 $P_{\theta_0}(p_\delta = 0)$		β_1 $P_{\theta_1}(p_\delta = 0)$
	Data are inconclusive		γ_0 $P_{\theta_0}(0 < p_\delta < 1)$		γ_1 $P_{\theta_1}(0 < p_\delta < 1)$
	Data are consistent with the null		ω_0 $P_{\theta_0}(p_\delta = 1)$		ω_1 $P_{\theta_1}(p_\delta = 1)$

ASA SGPV Short Course

Blume and Murray, 2022

Traditional p-value vs. SGPV

		Null is true		Alternative is true	
		Traditional p-value	SGPV	Traditional p-value	SGPV
Inference	Data are consistent with the alternative	Type I error $P_{\theta_0}(p \leq 0.05)$	β_0 $P_{\theta_0}(p_\delta = 0)$	Power or (1 - Type II error) $P_{\theta_1}(p \leq 0.05)$	β_1 $P_{\theta_1}(p_\delta = 0)$
	Data are inconclusive	1 - Type I error $P_{\theta_0}(p > 0.05)$	γ_0 $P_{\theta_0}(0 < p_\delta < 1)$	Type II error $P_{\theta_1}(p > 0.05)$	γ_1 $P_{\theta_1}(0 < p_\delta < 1)$
	Data are consistent with the null	NA	ω_0 $P_{\theta_0}(p_\delta = 1)$	NA	ω_1 $P_{\theta_1}(p_\delta = 1)$

ASA SGPV Short Course

Blume and Murray, 2022

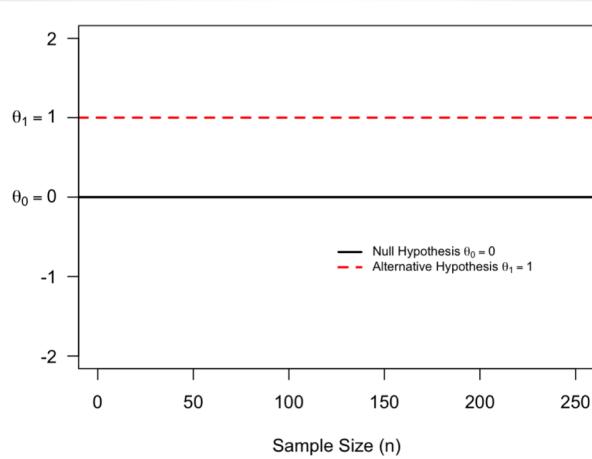
Traditional p-value vs. SGPV

		Null is true		Alternative is true	
		Traditional p-value	SGPV	Traditional p-value	SGPV
Inference	Data are consistent with the alternative	Type I error $P_{\theta_0}(p \leq 0.05)$	β_0 $P_{\theta_0}(p_\delta = 0)$	Power or (1 - Type II error) $P_{\theta_1}(p \leq 0.05)$	β_1 $P_{\theta_1}(p_\delta = 0)$
	Data are inconclusive	1 - Type I error $P_{\theta_0}(p > 0.05)$	γ_0 $P_{\theta_0}(0 < p_\delta < 1)$	Type II error $P_{\theta_1}(p > 0.05)$	γ_1 $P_{\theta_1}(0 < p_\delta < 1)$
	Data are consistent with the null	NA	ω_0 $P_{\theta_0}(p_\delta = 1)$	NA	ω_1 $P_{\theta_1}(p_\delta = 1)$

ASA SGPV Short Course

Blume and Murray, 2022

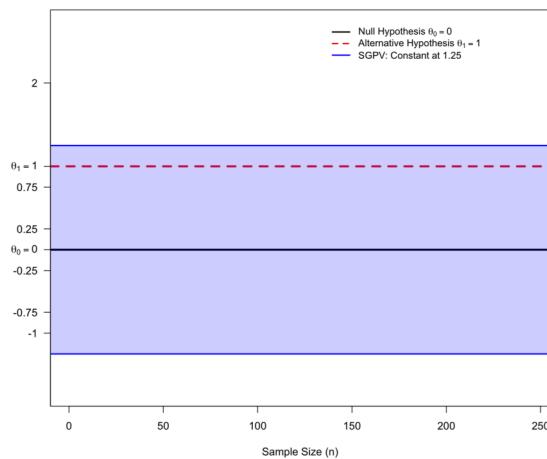
Standard Assumptions



ASA SGPV Short Course

Blume and Murray, 2022

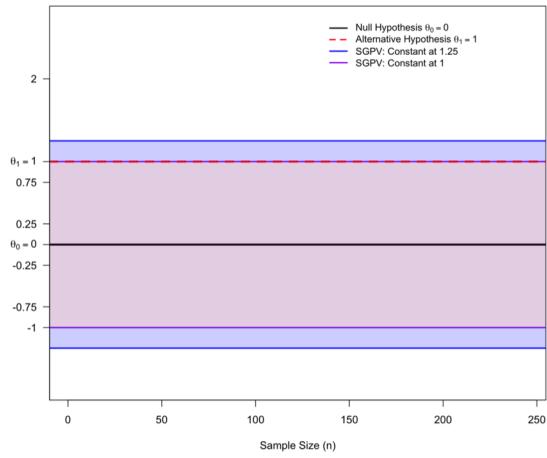
Example Indifference Zones



ASA SGPV Short Course

Blume and Murray, 2022

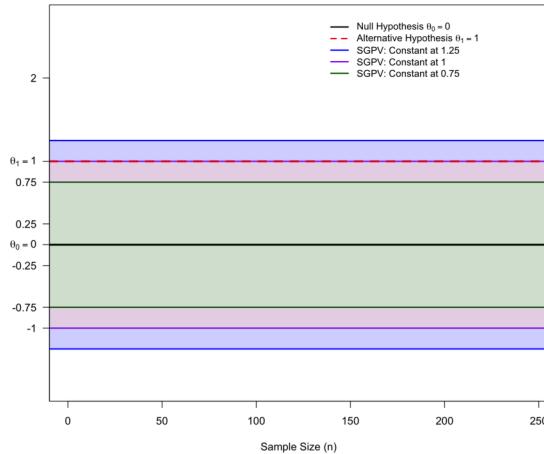
Example Indifference Zones



ASA SGPV Short Course

Blume and Murray, 2022

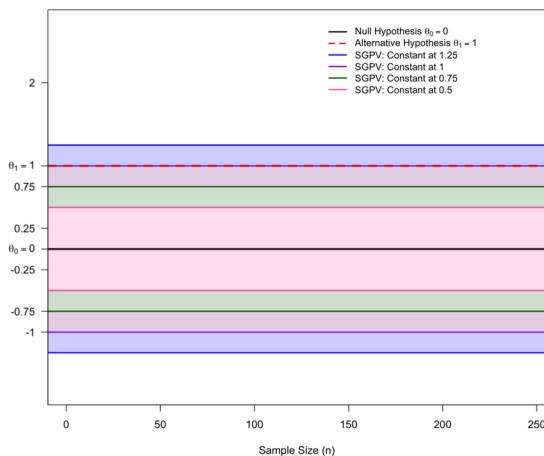
Example Indifference Zones



ASA SGPV Short Course

Blume and Murray, 2022

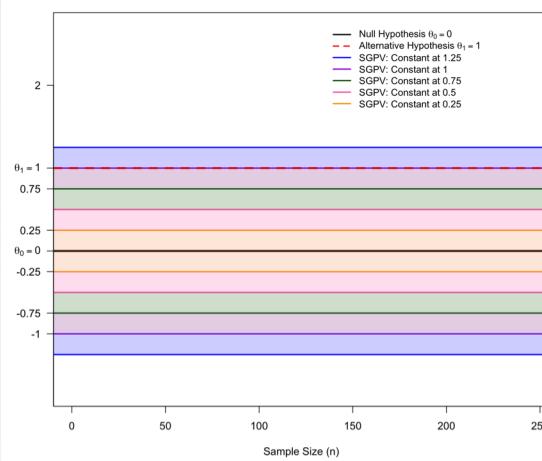
Example Indifference Zones



ASA SGPV Short Course

Blume and Murray, 2022

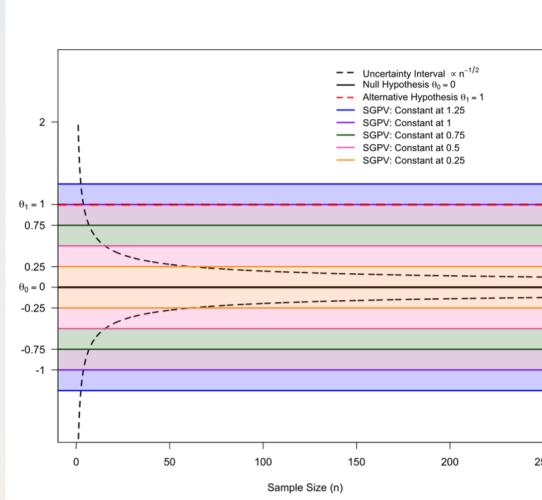
Example Indifference Zones



ASA SGPV Short Course

Blume and Murray, 2022

Now with CI centered at null



ASA SGPV Short Course

Blume and Murray, 2022

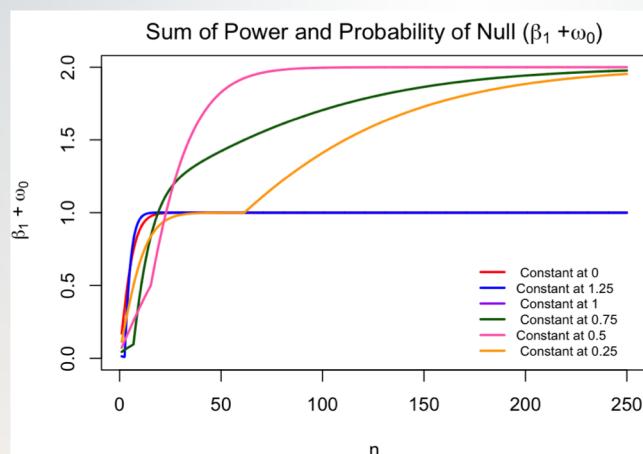
Traditional p-value vs. SGPV

		Null is true		Alternative is true	
		Traditional p-value	SGPV	Traditional p-value	SGPV
Inference	Data are consistent with the alternative	Type I error $P_{\theta_0}(p \leq 0.05)$	β_0 $P_{\theta_0}(p_\delta = 0)$	Power or (1 - Type II error) $P_{\theta_1}(p \leq 0.05)$	β_1 $P_{\theta_1}(p_\delta = 0)$
	Data are inconclusive	1 - Type I error $P_{\theta_0}(p > 0.05)$	γ_0 $P_{\theta_0}(0 < p_\delta < 1)$	Type II error $P_{\theta_1}(p > 0.05)$	γ_1 $P_{\theta_1}(0 < p_\delta < 1)$
	Data are consistent with the null	NA	ω_0 $P_{\theta_0}(p_\delta = 1)$	NA	ω_1 $P_{\theta_1}(p_\delta = 1)$

ASA SGPV Short Course

Blume and Murray, 2022

Indifference Zone Properties



- Maximize power and probability of true nulls
- Plot $\beta_1 + \omega_0$ over sample size
- The half a standard deviation between θ_1 and θ_0 gives the highest sum for reasonable sample sizes

ASA SGPV Short Course

Blume and Murray, 2022

Shrinking Indifference Zones

- Collaborator is uncertain but comfortable with a wider interval to begin
For example: $[-0.9, 0.9]$
- Limit shrinking to half distance between $(\theta^- \text{ or } \theta^+)$ and θ_0 $[-0.45, 0.45]$
- We can let the indifference zone shrink over sample size at different rates:

$$\delta_1(n) \propto \frac{\sqrt{V}}{n^{1/4}}$$

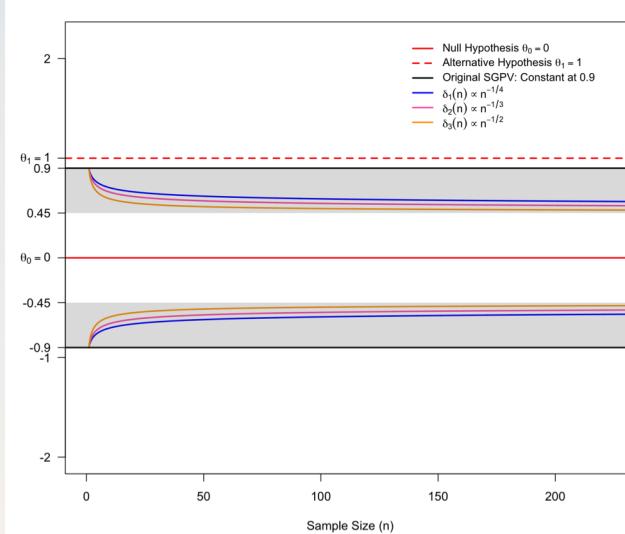
$$\delta_2(n) \propto \frac{\sqrt{V}}{n^{1/3}}$$

$$\delta_3(n) \propto \frac{\sqrt{V}}{n^{1/2}}$$

ASA SGPV Short Course

Blume and Murray, 2022

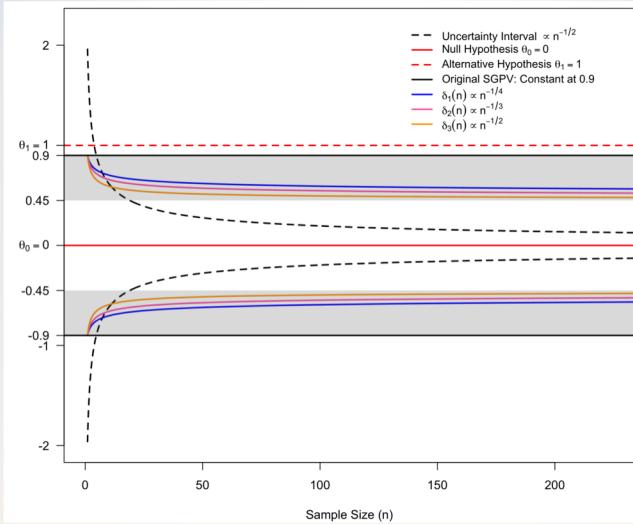
Shrinking Indifference Zones



ASA SGPV Short Course

Blume and Murray, 2022

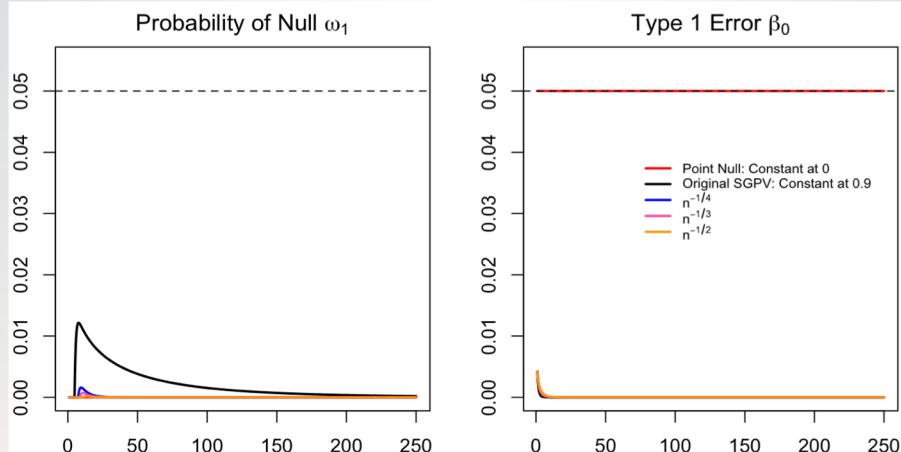
Now with CI centered at null



ASA SGPV Short Course

Blume and Murray, 2022

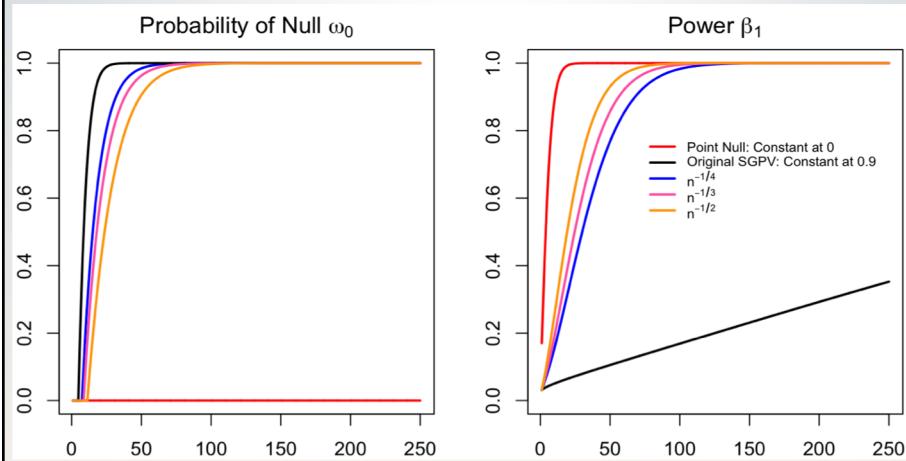
Incorrect probabilities (Error)



ASA SGPV Short Course

Blume and Murray, 2022

Correct probabilities (Power)



ASA SGPV Short Course

Blume and Murray, 2022

Recommendations for Collaborator Uncertainty

Collaborator Hypothesis	Suggested SGPV Analysis	Outcome
Confident in a null zone.	Use it!	Ideal
Confident in the alternative point closest to null.	Use the halfway point as the null zone.	Great
Uncertain of null zone.	Use the null zone that shrinks.	Good
Uncertain of alternative point.	Use the null zone that shrinks.	Ok
Cannot identify.	Do NOT use SGPV! (but can't ever measure evidence for null)	Poor

ASA SGPV Short Course

Blume and Murray, 2022

SGPV Study Planning Conclusions

- Goal is to balance statistical properties
- Results and inference will vary depending on the chosen indifference zone
- Promotes conversation between collaborator and statistician
 - Point null
 - Indifference zone or closest alternative point of interest
 - Sample size
 - If shrinking is necessary

ASA SGPV Short Course

Blume and Murray, 2022

Equivalence Tests

- Establish bioequivalence between data and an established equivalence range or interval null
- Example: A pharmaceutical company tests for drug approval by comparing new drug's performance to an approved drug's performance
- Uses an interval null or equivalence range
 - $H_0 = [\theta^-, \theta^+]$

ASA SGPV Short Course

Blume and Murray, 2022

TOST Test

- Most popular frequentist test is the Two One-Sided t-tests (TOST) (Schuirmann 1987)
 - Flips the null and alternative (be careful)
 - Uses the $(1-2\alpha)\%$ confidence interval (be careful)
- Tests are ordinary, one-sided, α -level t-tests
- If *both* one-sided tests reject, then conclude the evidence is contained in the equivalence range

ASA SGPV Short Course

Blume and Murray, 2022

TOST Test

- **Fail to reject the (original) null:** The confidence interval is outside of the indifference zone

$$H_0: (\theta < \theta^- \text{ or } \theta > \theta^+)$$
- **Reject the (original) null:** The confidence interval is contained within the indifference zone

$$H_1: (\theta \geq \theta^- \text{ and } \theta \leq \theta^+)$$
- Reported p -value is the p -value of largest magnitude from the two one-sided tests

$$p_T = \max\{p_{T_1}, p_{T_2}\}$$

ASA SGPV Short Course

Blume and Murray, 2022

TOST vs. SGPV comparison

		SGPV Outcomes		
		Consistent with the alternative (SGPV near 0)	Inconclusive (SGPV near $\frac{1}{2}$)	Consistent with the null (SGPV near 1)
Equivalence Tests Outcomes	Consistent with the alternative (p -value is unable to indicate this)	A	B	C
	Inconclusive (p -value is non-significant)	D	E	F
	Consistent with the null (p -value is significant)	H	I	J

ASA SGPV Short Course

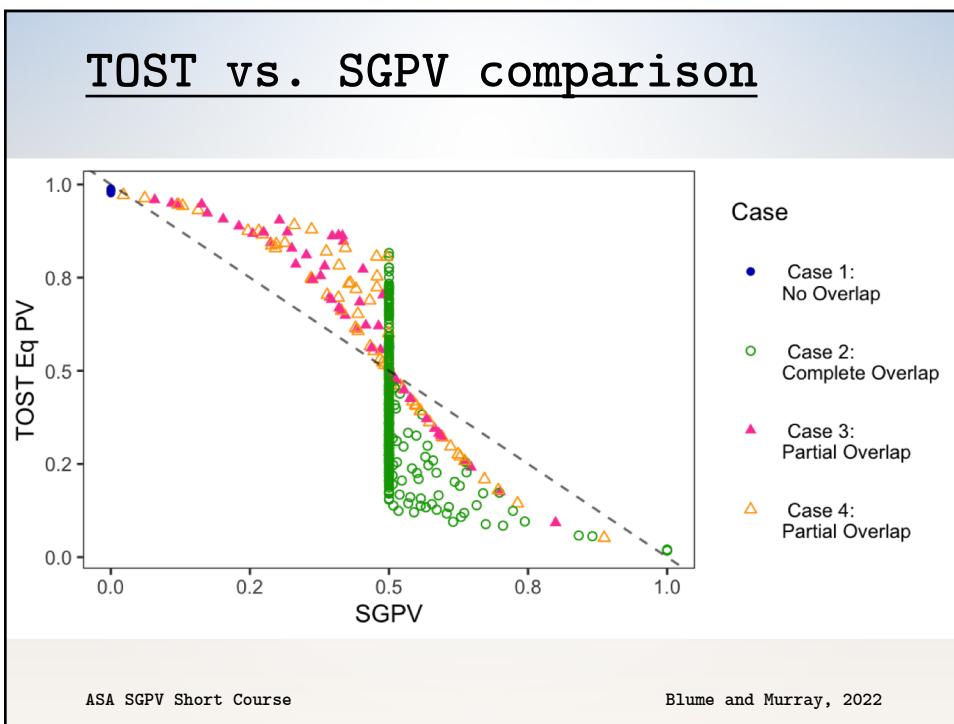
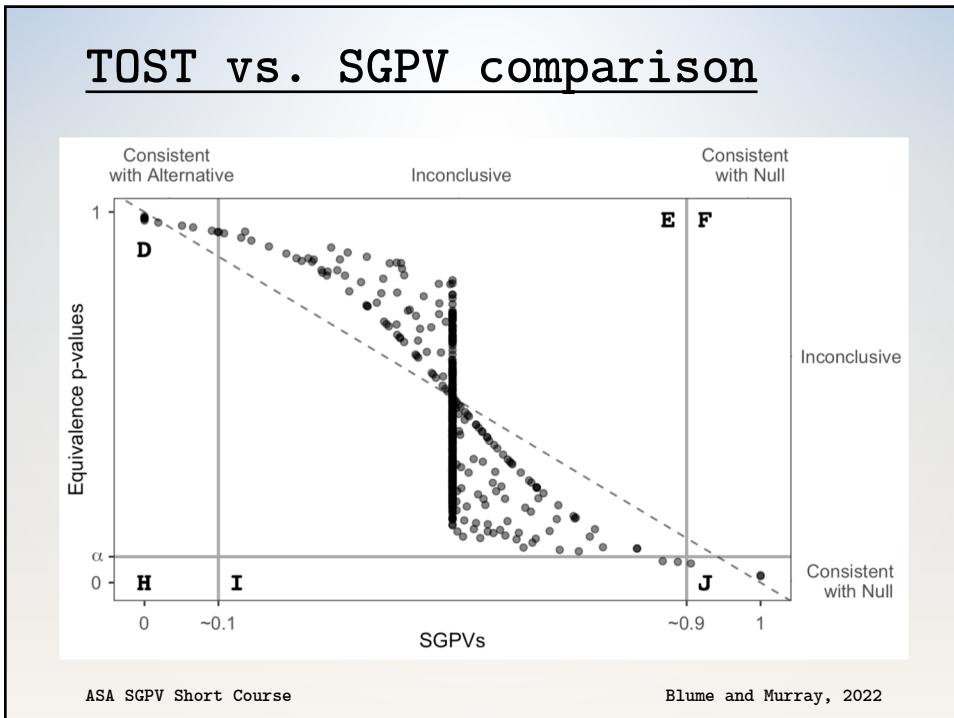
Blume and Murray, 2022

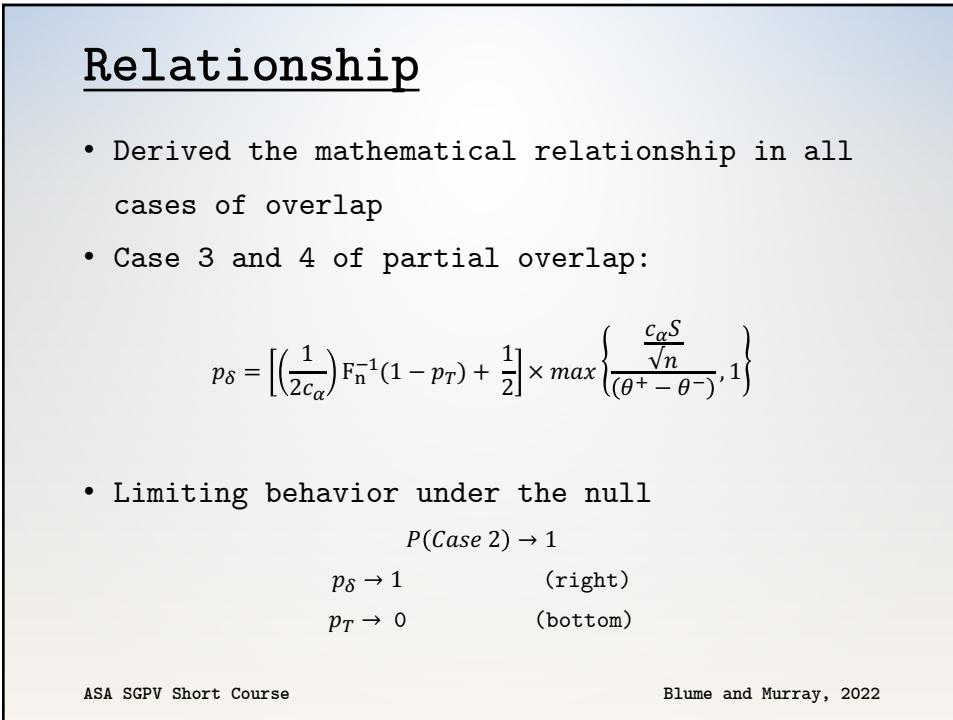
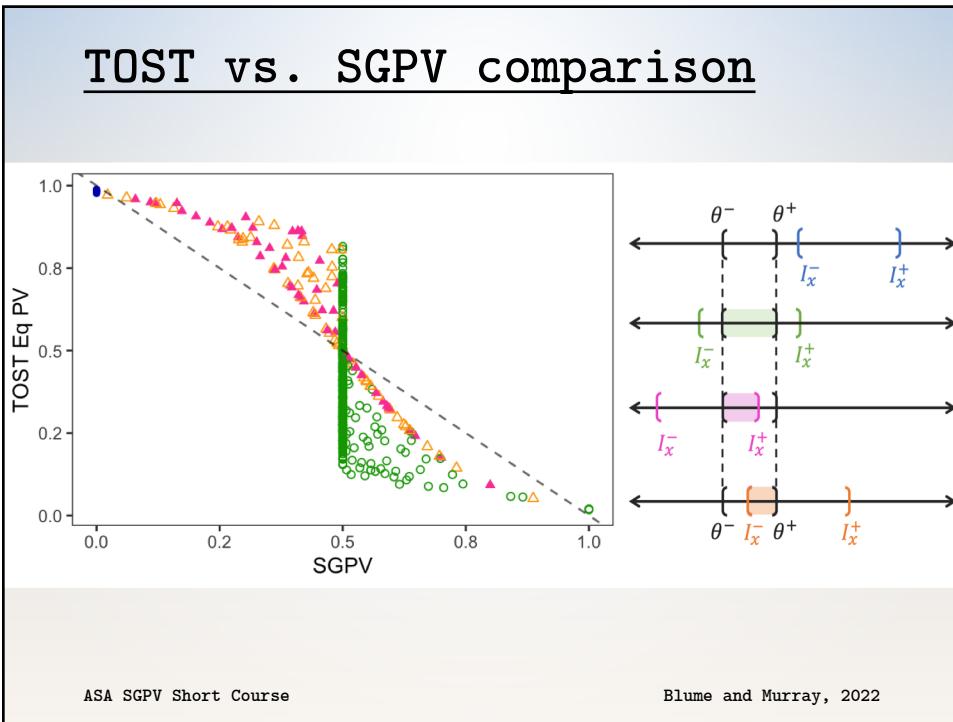
TOST v SGPV Simulation

- Simulate TOST and SGPV reported p -values
 - Data generated under the null, $N(0,1)$
 - Sample size of $n=6$
 - Yields 70% power for $\Delta = 1$ with 5% type 1 error
 - Indifference zone is $[\theta^-, \theta^+] = [-0.375, 0.375]$
 - Uncertainty interval is 95% confidence interval
 - 500 iterations (for illustration)
 - TOSTER R package ([Link](#))

ASA SGPV Short Course

Blume and Murray, 2022





TOST vs. SGPV comparison

TOST	SGPV
• 2 inference outcomes	• 3 inference outcomes
• Conclusions only about $(1 - 2\alpha)\%$ confidence interval	• Any uncertainty data interval can be used
• Type I Error is ultra-conservative (distribution of p_T is non-uniform)	• Type I error is accurately assessed (Tied to width of data interval)
• Not uniformly most powerful	• Indicates when data agree with null or alternative without additional testing
• No measure of overlap included in computation	• Includes overlap in reported p -value

ASA SGPV Short Course

Blume and Murray, 2022

Variable Selection

- SGPV variable selection (regression modeling) by Yi Zuo
 - Jeffrey's student
 - Yi's dissertation work January 2022
 - Currently working at Merck
 - R Package: ProSGPV

ASA SGPV Short Course

Blume and Murray, 2022

Variable Selection

- Traditional p-values do not perform well in variable selection
 - Tends to include trivial effects and results are sensitive to small modifications
- Second-generation p-values can improve variable selection using clinical significance
 - Superior support recovery, parameter estimation, and even prediction in certain scenarios, when compared to current standard procedures
 - Can accommodate continuous, binary, count, and time-to-event data
 - An R package made it easy to implement

ASA SGPV Short Course

Blume and Murray, 2022

Introduction

- Data are typically comprised of an outcome and features.
- A common scientific task is to separate the relevant features from the noise features.
- This task is called support recovery, which involves variable selection.
- We also want precise & unbiased parameter estimation, and good prediction

ASA SGPV Short Course

Blume and Murray, 2022

Current Approaches

- P-value based approaches
- l_0 -based approach
- l_1 -based approaches
- l_0 and l_1 hybrid approaches
- l_1 and l_2 hybrid approach
- Marginal correlation based approach

ASA SGPV Short Course

Blume and Murray, 2022

Drawbacks of current approaches

- All procedures (particularly, adaptive lasso, ISIS, SCAD, and MC+) are great, in theory.
- The actual variable selection results depend on tuning parameters that are hard to specify in practice.
- A prediction-optimal lasso selects all signals + noise variables. It is a good place to start with.

ASA SGPV Short Course

Blume and Murray, 2022

Second-generation p-values

- Second-generation p-values (SGPVs) were proposed in the high dimensional multiple testing context (Blume, D'Agostino McGowan, et al. 2018; Blume, Greevy, et al. 2019).
- SGPVs replace the point null hypothesis with a pre-specified interval null, which can be used to select effects that are clinically meaningful.

ASA SGPV Short Course

Blume and Murray, 2022

Second-generation p-values

- SGPV enjoys several appealing properties:
 - SGPV close to 0 indicates support for the alternative hypothesis; close to 1 indicates support for the null hypothesis; and near 1/2 is inconclusive.
 - It doesn't need a threshold in the interpretation.
 - It values clinically meaningful effects over traditionally statistically significant effects, which could be valuable to support recovery.

ASA SGPV Short Course

Blume and Murray, 2022

Penalized regression with Second-Generation P-Values (ProSGPV)

Algorithm 1 The ProSGPV algorithm

```

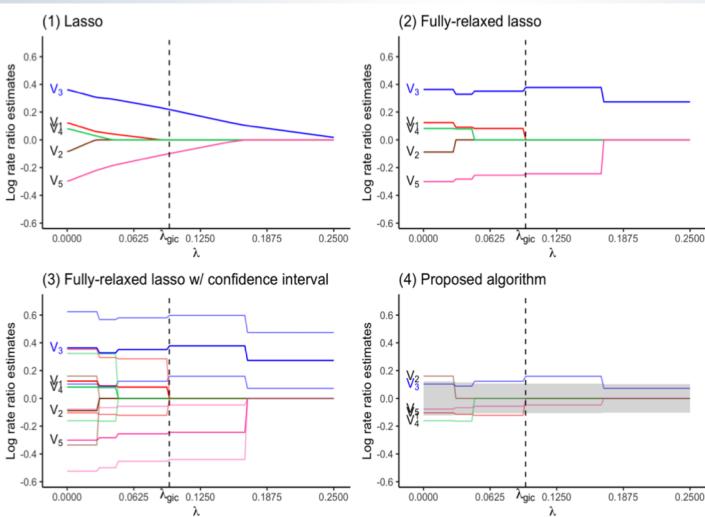
1: procedure PROSGPV( $\mathbf{X}$ ,  $\mathbf{Y}$ )
2:   Stage one: Find a candidate set
3:     Standardize explanatory variables
4:     Fit a lasso and evaluate it at  $\lambda_{\text{gic}}$ 
5:     Fit OLS/GLM/Cox models on the lasso active set
6:   Stage two: SGPV screening
7:     Extract the confidence intervals of all variables from the previous step
8:     Calculate the mean coefficient standard error  $\bar{SE}$ 
9:     Calculate the SGPV for each variable where  $I_j = \hat{\beta}_j \pm 1.96 \times SE_j$  and  $H_0 = [-\bar{SE}, \bar{SE}]$ 
10:    Keep variables with SGPV of zero
11:    Refit the OLS/GLM/Cox with selected variables
12: end procedure

```

ASA SGPV Short Course

Blume and Murray, 2022

ProSGPV Illustration



- Note that the null region in figure (4) is in grey. The value is $2 \times SE$ and it converges to 0 at \sqrt{n} rate.

ASA SGPV Short Course

Blume and Murray, 2022

Simulation

- Generate simulation data
- Run ProSGPV, lasso, BeSS, and ISIS on the data and record support recovery rate, parameter estimation mean absolute error (MAE), prediction root mean square error (RMSE) and area under the curve in a separate test set, and running time.
- Compare the performance of four algorithms over 1000 repetitions

ASA SGPV Short Course

Blume and Murray, 2022

Simulation Parameters

Table 1: Summary of parameters in simulation studies.

	Linear regression	Logistic regression	Poisson regression	Cox regression
n	100	32:320	40	80:800
p	100:1000	16	40:400	40
s	10	6	4	20
β_l	1	0.4	0.2	0.3
β_u	2	1.2	0.5	1
ρ	0.3	0.6	0.3	0.3
σ	2	2	2	2
ν	2			
Intercept t		0	2	0
Scale				2
Shape				1
Rate of censoring				0.2

ASA SGPV Short Course

Blume and Murray, 2022

Metrics

- Support recovery is defined as capturing the exact true support, not containing it.
- An estimate of MAE is the following where $\beta_{0,j}$ is the jth true coefficient

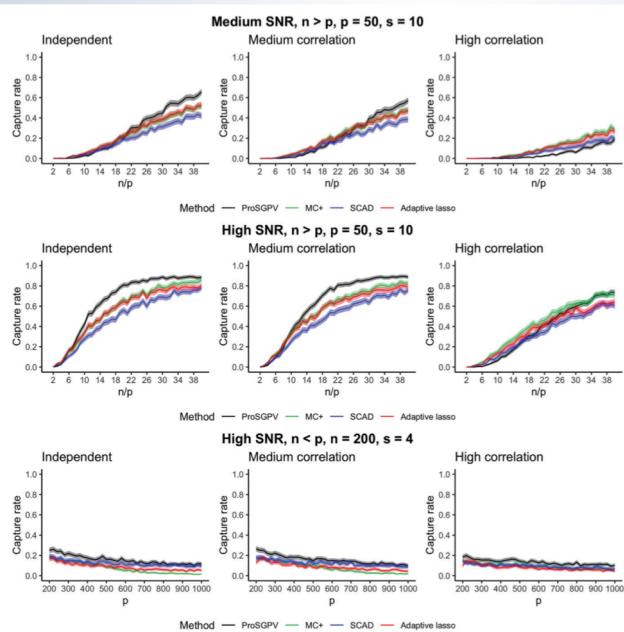
$$\frac{1}{p} \sum_{j=1}^p \| \hat{\beta}_j - \beta_{0,j} \|$$

- MSE is the usual mean squared error

ASA SGPV Short Course

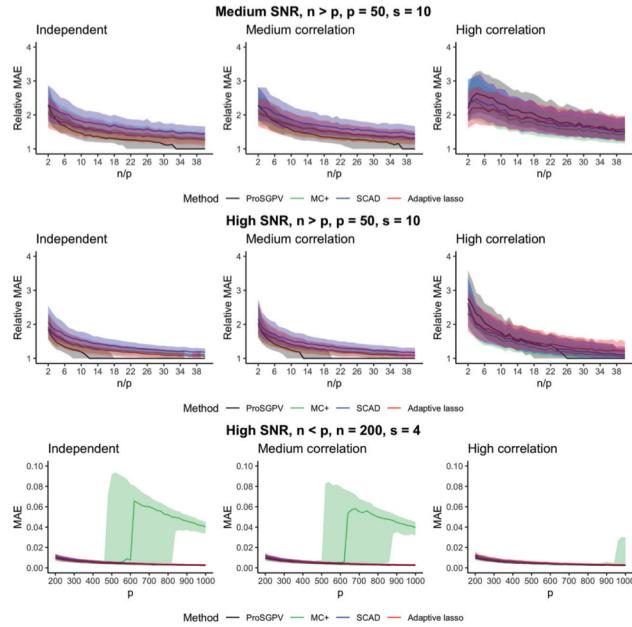
Blume and Murray, 2022

Capture Rate



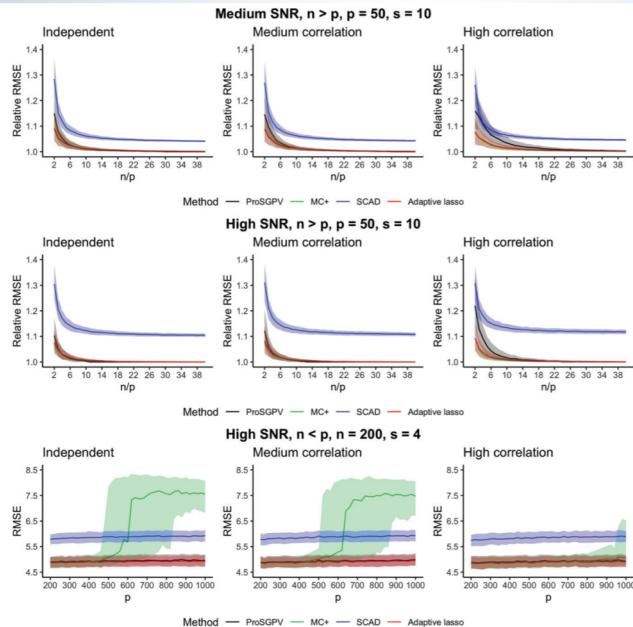
- Capture rate of the exact true model under combinations of autocorrelation level, signal-noise-ratios, and (n, p, s). In each panel, one algorithm has a colored solid line representing the average capture rate surrounded by the shaded 95% Wald interval over 1000 simulations.

Parameter Estimation Error

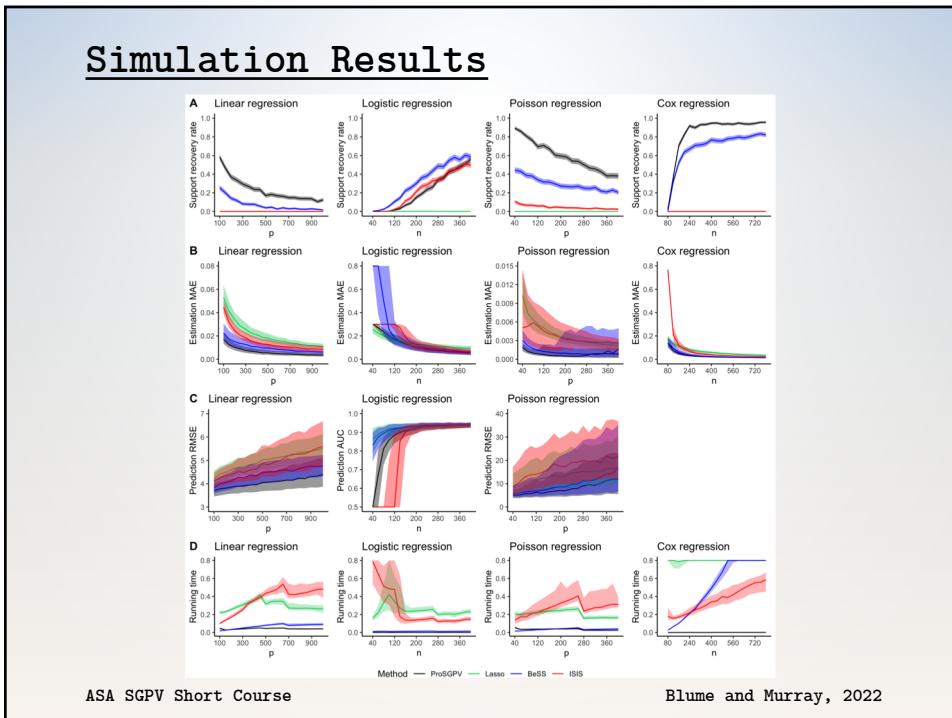
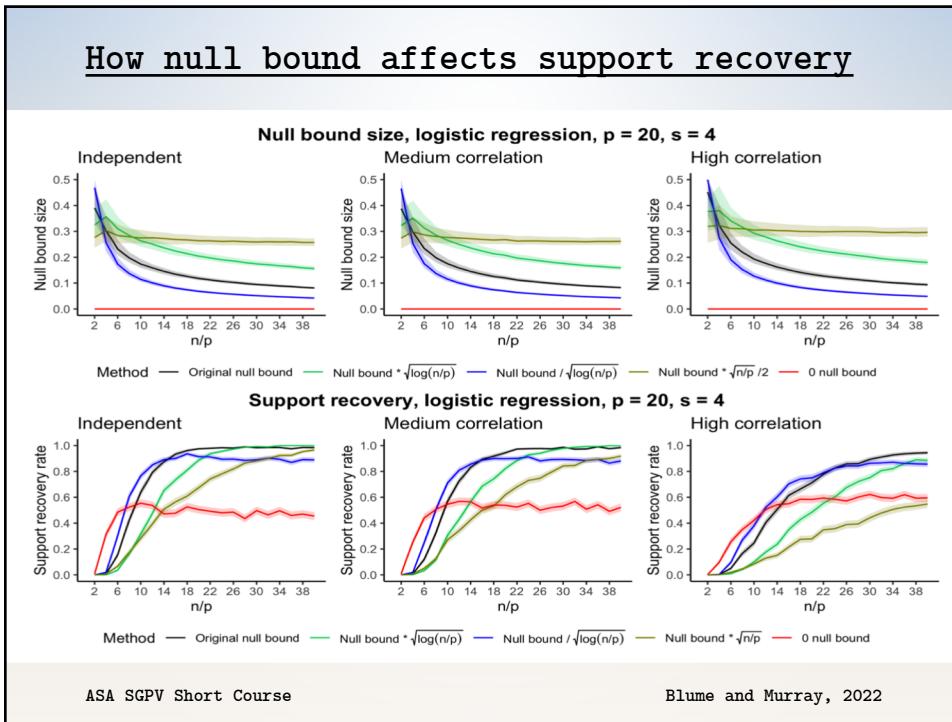


- Parameter estimation error of all algorithms under combinations of autocorrelation level, SNR, and (n, p, s) . In each panel, one algorithm has a colored solid line representing the median (relative) MAEs surrounded by the shaded first and third quartiles over 1000 simulations.

Prediction Accuracy



- Comparison of prediction accuracy of all algorithms under combinations of autocorrelation level, SNR, and (n, p, s) . Median (relative) root mean square errors are surrounded by their first and third quartiles over 1000 simulations.



Takeaways

- Traditional p-values do not perform well in variable selection
 - Tends to include trivial effects and results are sensitive to small modifications
- Second-generation p-values can improve variable selection using clinical significance
 - Superior support recovery, parameter estimation, and even prediction in certain scenarios, when compared to current standard procedures
 - Can accommodate continuous, binary, count, and time-to-event data
 - An R package made it easy to implement

ASA SGPV Short Course

Blume and Murray, 2022

Variable Selection with SGPVs

- ProSGPV GitHub
 - <https://github.com/zuoyi93/ProSGPV>
- Papers
 - <https://www.tandfonline.com/doi/full/10.1080/00031305.2021.1946150>
 - f1000research.com/articles/11-58
- Linear ProSGPV
 - <https://cran.r-project.org/web/packages/ProSGPV/vignettes/linear-vignette.html>
- GLM and Cox ProSGPV
 - <https://cran.r-project.org/web/packages/ProSGPV/vignettes/glm-cox-vignette.html>

ASA SGPV Short Course

Blume and Murray, 2022

Review of Topics

- Second-generation p -value framework, definition, and examples
- The SGPV achieves the inferential properties that many scientists hope, or believe, are attributes of the classic p -value.
- Statistical Properties of SGPs
- False Discovery Rates
- Study Planning
- Comparison to Equivalence Tests: Two One-Sided Tests (TOST)
- SGPV Variable Selection
- Coding Examples

Blume and Murray, 2022

Acknowledgements

- Collaborators
 - William D. Dupont
 - Robert A. Greevy
 - Lucy D'Agostino McGowan
 - Jonathan Chipman
 - Valerie Welty
 - Lisa Lin
 - Jeffrey R. Smith
 - Yi Zuo
 - Thomas G. Stewart
 - Vanderbilt SEDS Lab
- Website
 - www.statisticalevidence.com

Questions?