Covariate Adjustment

Michael Martens, PhD

November 3, 2022

Covariate Adjustment

- In a clinical trial, primary interest is the treatment's effect on outcome
- Other covariates may also affect outcome (e.g. age, genetics)
- Why adjust for covariates?
  - Increased precision of estimated treatment effect
  - Imbalances in patient characteristics can occur even with randomization, so could confound analysis if not adjusted
  - Learn about effects of covariates

Covariate Adjustment

- How do we adjust?
  - Stratification:
    - Divide subjects into smaller, more homogeneous subgroups (i.e. strata)
    - Compare treatments separately for each stratum, then combine estimates of the treatment effect across strata
  - Regression modeling: Covariates used as predictors in a regression model

Covariate Adjustment

- Which covariates do we adjust for?
  - Baseline demographics, disease, and/or treatment-related variables
  - Should be related to outcome; can assess this based on prior research and expert knowledge
  - Best to specify them upfront in the study protocol
  - Often the same ones used for stratified randomization if employed

Covariate Adjustment: Normal Outcomes

- Suppose Y is a continuous outcome of interest and μ_i is the mean of Y under treatment i, $i = 1, 2$
- Suppose we want to test $H_0 : \mu_1 = \mu_2$ while adjusting for a baseline covariate X
- **Analysis of Covariance (ANCOVA)** model specifies that observation Y_{ij} for patient j in treatment group i is

$$Y_{ij} = \mu_i + \gamma b_{ij} + \epsilon_{ij}, \text{ where } \epsilon_{ij} \sim N(0, \sigma^2_\epsilon)$$

Covariate Adjustment: Normal Outcomes

- Model assumptions:
  - Linear relationship between response and baseline covariates
  - Equal treatment effects at all covariate values (no interactions)
  - Independent observations
  - Equal variance of error terms for all observations
- Estimate parameters using Least Squares approach
- SAS PROC GLM can fit model

Example

- A trial by Hommel et al. (1986) randomized insulin-dependent patients with diabetic nephropathy to receive either Captopril (drug to reduce BP) or placebo
- Systolic BP was measured before randomization and one week post treatment
- Primary outcome: Systolic BP at one week post treatment

		Pre		Post	
	n	Mean	SD	Mean	SD
Captopril	9	148.0	11.4	135.33	8.43
Placebo	7	146.6	12.3	141.86	6.94

Example

- A t-test comparing post treatment measurements yields $t = 1.65$, $p = 0.12$
- Baseline BP appears balanced ($p = 0.81$)
  - Not informative, because randomization is expected to produce balance
- Baseline BP is expected to be related to BP after treatment
- In a given trial, baseline BP may be imbalanced unless stratified randomization is used
- Using ANCOVA to adjust for baseline BP, $p = 0.031$, finding a reduction of systolic BP under Captopril treatment

Example: ANCOVA in SAS

```
proc glm data=captopril;
  class trt(ref='0');
  model post=trt baseline /solution clparm;
  lsmeans trt;
run;
```

Example: ANCOVA in SAS

Parameter		Estimate		Standard Error	t Value	Pr > |t|
Intercept		74.75098814	B	19.59679310	3.81	0.0021
trt	1	-7.17786561	B	2.96364143	-2.42	0.0308
trt	0	0.00000000	B	.	.	.
baseline		0.45783926		0.13284206	3.45	0.0043

Parameter		95% Confidence Limits	
Intercept		32.41469056	117.08728573
trt	1	-13.58042367	-0.77530755
trt	0	.	.
baseline		0.17085144	0.74482709

The GLM Procedure
Least Squares Means

trt	post LSMEAN
1	135.047184
0	142.225049

Sample Size / Power Calculation for ANCOVA

- Suppose patients are randomized to treatment with $n_1 = rn_2$
- Want to detect a treatment effect $\delta = \mu_1 - \mu_2$ with $1 - \beta$ power using an α level test
- Can show that an appropriate sample size is

$$n_2 = \frac{(r+1)}{r} \frac{(z_{1-\alpha/2} + z_{1-\beta})^2 \sigma_\epsilon^2}{\delta^2},$$

where σ_ϵ^2 is the residual variance of the ANCOVA model

Example

- Trial Design: Randomized, placebo-controlled trial to assess effectiveness of zanamivir, a new flu treatment
- Primary outcome: # of days to alleviation of symptoms
- Age may affect symptom duration, so model will adjust for it (≤ 65 vs. > 65 years)
- Assume residual standard deviation is $\sigma_\epsilon = 2.5$, and investigators want to detect a reduction of 1 day in the mean time until alleviation of symptoms ($\delta = -1$)
- Using a 5% two-sided significance level, power of 90%, note that $z_{1-\alpha/2} = 1.96$ and $z_{1-\beta} = 1.28$. Then

$$\frac{2(1.96 + 1.28)^2(2.5)^2}{(-1)^2} = 131.2 \approx 132 \text{ per group}$$

Binary Data

- With a binary outcome, the treatment effect is often described using the odds ratio:
  - **Odds** $= p/(1-p)$, where p is the probability of success
  - **Odds Ratio** for treatment 1 vs. 2:

$$OR = \frac{p_1/(1-p_1)}{p_2/(1-p_2)},$$

    where p_i is the probability of success for treatment i
  - Can estimate odds and OR by plugging in sample proportions

Example

- Trial comparing two treatments for psoriasis
- PUVA therapy: drug called methoxsalen, followed by exposure to UVA (longer wavelength UV rays)
- TL-01 therapy: new lamp for administering shorter wavelength UV-B radiation
- Primary outcome: Clearance of psoriasis by end of study

	Cleared	Did not clear	Total
PUVA	41	8	49
TL-01	32	19	51
Total	73	27	100

Example

- Odds of clearance for PUVA $= 41/8 = 5.125$
- Odds of clearance for TL-01 $= 32/19 = 1.684$
- Estimated odds ratio for clearance for PUVA group vs. TL-01 group:

$$\widehat{OR} = \frac{5.125}{1.684} = 3.043$$

- Odds of clearance are 3 times higher for patients undergoing PUVA treatment compared to TL-01 treatment

Confidence Interval for Odds Ratio

- Assume we have the following binary response data:

	Success	Failure
Treatment	a	b
Control	c	d

- Construct CI based on log transformation: For large n,

$$SE(\log \widehat{OR}) \approx \sqrt{\frac{1}{a} + \frac{1}{b} + \frac{1}{c} + \frac{1}{d}}$$

- Then a 95% confidence interval for $\log OR$ is

$$\log \widehat{OR} \pm 1.96 \cdot SE(\log \widehat{OR}) = [L, U]$$

- A 95% confidence interval for OR is $[e^L, e^U]$

Example

- Standard Error:

$$SE(\log \widehat{OR}) = \sqrt{\frac{1}{32} + \frac{1}{41} + \frac{1}{19} + \frac{1}{8}} = 0.483$$

- 95% confidence interval for $\log OR$:

$$\log(3.043) \pm 1.96(0.483) = [0.166, 2.059]$$

- 95% confidence interval for OR:

$$[e^{0.166}, e^{2.059}] = [1.181, 7.842]$$

Stratified Analysis of Binary Outcomes

- Stratification involves assessing treatment effect separately in each stratum and then combining effects
- Pool estimates of odds ratio across strata's 2x2 tables
- Assumption: Odds ratio is constant across strata (no interaction)
- Cochran-Mantel-Haenszel test: Let r_{ik} denote the number of successes in group i for stratum k

	Success	Failure	Total
Experimental	r_{1k}	$n_{1k} - r_{1k}$	n_{1k}
Control	r_{2k}	$n_{2k} - r_{2k}$	n_{2k}
Total	r_k	$n_k - r_k$	n_k

Stratified Analysis of Binary Outcomes

- Expected value and variance of r_{1k} are

$$E(r_{1k}) = e_{1k} = r_k n_{1k}/n_k \text{ and}$$

$$Var(r_{1k}) = V_{1k} = \frac{n_{1k} n_{2k} r_k (n_k - r_k)}{n_k^2 (n_k - 1)}$$

- Test statistic: $U = \sum_k (r_{1k} - e_{1k})$
- $Var(U) = V = \sum_k V_{1k}$ since strata are independent
- Under $H_0 : OR = 1$, $E(U) = 0$ so that for large samples,

$$C = \frac{U^2}{V} \sim \chi_1^2$$

Stratified Analysis of Binary Outcomes

- Combined estimate of the odds ratio:

$$\widehat{OR}_{MH} = \frac{\sum_k r_{1k}(n_{2k} - r_{2k})/n_k}{\sum_k r_{2k}(n_{1k} - r_{1k})/n_k}$$

- Confidence interval methods available in SAS

Example

Stratified analysis of PUVA vs. TL-01 data, based on plaque size

	Small plaques			Large plaques		
	Cleared	Not	Total	Cleared	Not	Total
PUVA	25	3	28	16	5	21
TL-01	23	6	29	9	13	22
Total	48	9	57	25	18	43

Example

- Small plaques: $r_{11} = 25,\ e_{11} = 28 \cdot 48/57 = 23.579,$

$$V_{11} = \frac{29 \cdot 28 \cdot 48 \cdot 9}{57^2 \cdot 56} = 1.928$$

- Large plaques: $r_{12} = 16,\ e_{12} = 21 \cdot 25/43 = 12.209,$

$$V_{12} = \frac{22 \cdot 21 \cdot 25 \cdot 18}{43^2 \cdot 42} = 2.677$$

- Cochran-Mantel-Haenszel statistic:

$$C = \frac{[(25 - 23.579) + (16 - 12.209)]^2}{1.928 + 2.677} = 5.899,$$

so that $p = 0.015$

Example

- Stratified estimate of odds ratio

$$\widehat{OR}_{MH} = \frac{25 \cdot 6/57 + 16 \cdot 13/43}{23 \cdot 3/57 + 9 \cdot 5/43} = 3.309,$$

  with 95% CI of $[1.236, 8.860]$

- Contrast this with unstratified result: $\widehat{OR} = 3.043$, 95% CI $[1.181, 7.842]$, $p = 0.018$

Example: CMH Test in SAS

```
data psoriasis; input trt $ plaque $ success cnt; cards;
P S 1 25
P S 0 3
P L 1 16
P L 0 5
T S 1 23
T S 0 6
T L 1 9
T L 0 13
;run;

proc freq data=psoriasis order=data;
  weight cnt;
  tables plaque*trt*success/relrisk cmh;
run;
```

Example: CMH Test in SAS

The FREQ Procedure
Summary Statistics for trt by success
Controlling for plaque

Cochran-Mantel-Haenszel Statistics (Based on Table Scores)

Statistic	Alternative Hypothesis	DF	Value	Prob
1	Nonzero Correlation	1	5.8983	0.0152

Common Odds Ratio and Relative Risks

Statistic	Method	Value	95% Confidence Limits	
Odds Ratio	Mantel-Haenszel	3.3091	1.2359	8.8598

Regression Model-based Adjustment for Covariates

- **Logistic Regression**:

$$\text{logit } P(Y=1|Z,\mathbf{X}) = \log \frac{P(Y=1|Z,\mathbf{X})}{1-P(Y=1|Z,\mathbf{X})} = \beta_0 + \beta Z + \boldsymbol{\gamma}'\mathbf{X},$$

  where $Z = 1$ for experimental, 0 if control treatment,
  $\mathbf{X}$ is the vector of baseline covariates

- For a patient assigned to experimental treatment,

$$\text{logit } P(Y=1|Z=1,\mathbf{X}) = \beta_0 + \beta + \boldsymbol{\gamma}'\mathbf{X}$$

- For a patient assigned to control treatment,

$$\text{logit } P(Y=1|Z=0,\mathbf{X}) = \beta_0 + \boldsymbol{\gamma}'\mathbf{X}$$

Model-based Adjustment for Covariates

- Then the log odds ratio of treatment adjusted for **X** is

$$\log OR_{trt} = \text{logit } P(Y = 1|Z = 1, \mathbf{X}) - \text{logit } P(Y = 1|Z = 0, \mathbf{X}) = \beta$$

- Interpretation: $\exp(\beta)$ is the odds ratio for treatment vs. control for two patients with the same covariates, also called **Adjusted Odds Ratio**
- Logistic regression model is fit by maximum likelihood (more in Methods 2 course), implemented in SAS PROC LOGISTIC

Example: Logistic Regression in SAS

```
proc logistic data=psoriasis descending;
  weight cnt;
  class plaque(ref='S') trt(ref='T') / param=ref;
  model success=trt plaque;
run;
```

Example: Logistic Regression in SAS

Class	Value	Design Variables
plaque	L	1
	S	0
trt	P	1
	T	0

Analysis of Maximum Likelihood Estimates

Parameter		DF	Estimate	Standard Error	Wald Chi-Square	Pr > ChiSq
Intercept		1	1.2000	0.3994	9.0271	0.0027
trt	P	1	1.2194	0.5098	5.7200	0.0168
plaque	L	1	-1.4351	0.4975	8.3203	0.0039

Odds Ratio Estimates

Effect		Point Estimate	95% Wald Confidence Limits	
trt	P vs T	3.385	1.246	9.195
plaque	L vs S	0.238	0.090	0.631

Sample Size / Power Calculation for Logistic Regression

- Suppose patients are randomized to treatments with $n_1 = rn_2$
- Investigators want to detect a change in response rate from π_2 under control to π_1 under experimental therapy
- Equivalently, want to detect the log odds ratio for treatment of $\beta = \delta$ corresponding to this difference
- Two approaches for sample size calculation:
  - Use direct sample size formula for logistic model
  - Use sample size for unadjusted two-sample test of proportions as an approximation

Sample Size / Power Calculation for Logistic Regression: Direct Approach

- From maximum likelihood theory, it can be shown that for large samples, $\sqrt{n}(\hat{\beta} - \beta) \sim N(0, f_0^{-1})$ where

$$f_0 = E_{H_0}[Var(Y|Z, \mathbf{X})] = E\left\{ \frac{exp(\beta_0 + \boldsymbol{\gamma}^T \mathbf{X})}{[1 + \exp(\beta_0 + \boldsymbol{\gamma}^T \mathbf{X})]^2} \right\}$$

- Can be shown that an appropriate sample size is

$$n_2 = \frac{(r+1)}{r} \frac{(z_{\alpha/2} + z_\beta)^2}{f_0 \delta^2}$$

- Advantage: Formula should be accurate under large sample properties of logistic model
- Drawback: Need to specify β_0, $\boldsymbol{\gamma}$, and distribution of $\mathbf{X}$ to get f_0; historical data could be used to estimate these

Sample Size / Power Calculation for Logistic Regression: Two-sample Approximate Approach

- In theory, covariate adjustment should boost power compared to an unadjusted comparison of proportions
- A sample size calculation based on a two-sample test should provide a sufficient sample size, e.g.

$$n_2 = \frac{(r+1)(z_{\alpha/2} + z_{\beta})^2}{4r(\sin^{-1}\sqrt{\pi_1} - \sin^{-1}\sqrt{\pi_2})^2}$$

- Advantage: Simpler calculation than direct approach
- Drawback: Formula may be inaccurate if covariate effects are not weak

Example

- A control treatment is expected to have a 35% response rate
- A clinically relevant difference in the response rate is determined to be a 10% improvement (absolute)
- Disease stage (early vs. advanced) also affects response rate, so want to adjust for it
- Assume a randomized clinical trial with equal sample sizes
- How large a sample size is needed to have 90% power to detect this difference using a two-sided test at a 5% level?

Example: Direct Approach

- Targeted treatment OR is $\frac{0.45/0.55}{0.35/0.65} = 1.519$ so that $\delta = 0.418$
- Disease stage X is binary, 50% chance of early
- Expect stage OR of 2 (early vs. advanced), so $\gamma = \log(2)$
- Under null, expect $P(Y = 1) = 35\%$; set $\beta_0 = \text{logit}(35\%) = -0.619$
- Estimate $f_0 = 0.239$ using specifications for X, β_0, γ
- The sample size needed is

$$n = \frac{2(1.96 + 1.28)^2}{0.239 \cdot 0.418^2} = 502.77 \approx 503 \text{ per group}$$

- The total sample size would be 1006 across both groups

Example: Two-Sample Approximate Approach

- The sample size needed is

$$\begin{aligned} n &= \frac{(1.96 + 1.28)^2}{2(\sin^{-1}\sqrt{0.45} - \sin^{-1}\sqrt{0.35})^2} \\ &= 501.9 \approx 502 \text{ per group} \end{aligned}$$

- The total sample size would be 1004 across both groups

Example: Direct Sample Size Formula Approach in SAS

```
proc power;
logistic alpha=0.05
vardist('treatment') = binomial(0.5,1)
vardist('diseaserisk') = binomial(0.5,1)
testpredictor = "treatment"
covariates = "diseaserisk"
intercept = -0.619
testoddsratio = 1.519
covoddsratios = 2.0
ntotal=.
power=0.9;
run;
```

Example: Direct Sample Size Formula Approach in SAS

The POWER Procedure
Likelihood Ratio Chi-square Test for One Predictor

Fixed Scenario Elements	
Method	Shieh-O'Brien approximation
Alpha	0.05
Intercept	-0.619
Test Predictor	treatment
Odds Ratio for Test Predictor	1.519
Covariates	diseaserisk
Covariate Odds Ratios	2
Nominal Power	0.9
Total Number of Bins	4

Computed N Total	
Actual Power	N Total
0.900	997

Survival Analysis: Adjusting for Covariates

- For a stratified analysis of survival times, we want to compare survival curves between arms within each stratum
- Let $S_{ki}(t)$ be the survival curve for patients in stratum k receiving treatment i
- Null hypothesis is that $H_0 : S_{k1}(t) = S_{k2}(t)$ for all t, k
- Log-rank test can be adapted for testing this

Survival Analysis: Adjusting for Covariates

- Stratified log-rank test: Data are divided into k strata
- Similar as with unstratified log-rank test, we construct 2 by 2 tables for each stratum k, failure time t_j:

Group	# failed	# survived	# at risk
1	d_{k1j}	$n_{k1j} - d_{k1j}$	n_{k1j}
2	d_{k2j}	$n_{k2j} - d_{k2j}$	n_{k2j}
Total	d_{kj}	$n_{kj} - d_{kj}$	n_{kj}

- Assumption: Failure and censoring times are independent

Survival Analysis: Adjusting for Covariates

- Under H_0 the expected number of failures in group 1 should be

$$e_{k1j} = n_{k1j} \cdot (d_{kj}/n_{kj})$$

- Stratified log-rank statistic is $U = \sum_{j,k}(d_{k1j} - e_{k1j})$ with variance $V = \sum_{j,k} V_{jk}$, where

$$V_{jk} = \frac{n_{k1j} n_{k2j} d_{kj}(n_{kj} - d_{kj})}{n_{kj}^2 (n_{kj} - 1)}$$

- Under H_0 and for large samples,

$$C = \frac{U^2}{V} \sim \chi_1^2$$

Example

- PUVA vs. TL-01: Outcome = # of visits until clearance of psoriasis
- Covariate: small vs. large plaque size
- Stratum 1: Large plaque size: $U_1 = 7.0673,\ V_1 = 5.3913$
- Stratum 2: Small plaque size: $U_2 = 10.9200,\ V_2 = 8.7736$
- Stratified log-rank test:

$$C = \frac{U^2}{V} = \frac{(U_1 + U_2)^2}{V_1 + V_2} = \frac{(7.0673 + 10.9200)^2}{5.3913 + 8.7736} = 22.8412$$

with $p < 0.001$

Example: Stratified Log-rank Test in SAS

```
data a;
  set d.puvavstl01trial;
run;
proc lifetest data=a notable;
  time visits*clearance(0);
  strata plaque_size/group=group;
run;
```

Example: Stratified Log-rank Test in SAS

Stratified Test of Equality over Group

Test	Chi-Square	DF	Pr > Chi-Square
Log-Rank	22.8418	1	<.0001

Regression Model-based Covariate Adjustment for Survival Analysis

- Regression models can be used for covariate-adjusted comparison of survival outcomes
- Because of the direct relationship of survival and hazard functions, the hazard function is usually modeled
- Let $\lambda(t|Z, \mathbf{X})$ be the hazard function for a patient with $Z = \mathsf{I}(\text{patient assigned to experimental treatment})$ and covariates $\mathbf{X}$

Cox Proportional Hazards Model

- **Cox Proportional Hazards Regression Model**:

$$\lambda(t|Z, \mathbf{X}) = \lambda_0(t) \exp(\beta Z + \boldsymbol{\gamma}' \mathbf{X})$$

  where $\lambda_0(t)$ is a baseline hazard function
- Consider two patients with identical covariates $\mathbf{X}$, where patient 1 is assigned to experimental and patient 2 to control
- The **Adjusted Hazard Ratio** for these patients is

$$HR_{trt} = \frac{\lambda(t|Z = 1, \mathbf{X})}{\lambda(t|Z = 0, \mathbf{X})} = \frac{\lambda_0(t) \exp(\beta + \boldsymbol{\gamma}' \mathbf{X})}{\lambda_0(t) \exp(\boldsymbol{\gamma}' \mathbf{X})} = \exp(\beta)$$

Cox Proportional Hazards Model

- Assumptions:
  - Effects of treatment and covariates are constant over time
  - Effects of treatment and covariates on hazard are multiplicative
  - Censoring and failure times are independent given treatment and covariates
- Estimation performed using partial likelihood (more in Applied Survival Analysis course)

Stratified Cox Model

- If effects on some variables are not proportional, we can use separate baseline hazard rates for their strata to adjust
- Let hazard function be $\lambda(t|Z, \text{Stratum } k)$ for patient in stratum k and with indicator $Z = \text{I(patient assigned to experimental treatment)}$
- **Stratified Cox Proportional Hazards Model**:

$$\lambda(t|Z, \text{Stratum } k) = \lambda_{k0}(t) \exp(\beta Z)$$

where the $\lambda_{k0}(t)$ are stratum-specific baseline hazard functions

Stratified Cox Model

- Consider two patients in the same stratum, where patient 1 is assigned to experimental and patient 2 assigned to control
- The **Adjusted Hazard Ratio** for these patients is

$$HR_{trt} = \frac{\lambda(t|Z=1, \text{Stratum=k})}{\lambda(t|Z=0, \text{Stratum=k})} = \frac{\lambda_{k0}(t)\exp(\beta)}{\lambda_{k0}(t)\exp(0)} = \exp(\beta)$$

- Assumptions:
  - Effect of treatment is constant over time
  - Effect of treatment on hazard is multiplicative
  - Censoring and failure times are independent given treatment and stratum

Example

- Cox regression: HR of clearance for PUVA vs. TL-01 is 3.08 with 95% CI $[1.89, 5.00]$, $p < 0.001$
- Stratified Cox model: HR of clearance for PUVA vs. TL-01 is 3.02 with 95% CI $[1.86, 4.90]$, $p < 0.001$

Example: Cox Model and Stratified Cox Model in SAS

```
/* Ordinary Cox model */
proc phreg data=a;
  class plaque_size group;
  model visits*clearance(0)=plaque_size group/rl;
run;

/* Stratified Cox model */
proc phreg data=a;
  class plaque_size group;
  model visits*clearance(0)= group/rl;
  strata plaque_size;
run;
```

Example: Cox Model and Stratified Cox Model in SAS

```
/* Cox Model Results */
                  Analysis of Maximum Likelihood Estimates
                                  Parameter      Standard
Parameter              DF        Estimate         Error    Chi-Square    Pr > ChiSq
PLAQUE_SIZE Large       1       -0.94532       0.25237       14.0310        0.0002
GROUP       PUVA        1        1.12429       0.24780       20.5849        <.0001

                        Hazard      95% Hazard Ratio
Parameter                Ratio      Confidence Limits     Label
PLAQUE_SIZE Large        0.389       0.237       0.637    PLAQUE_SIZE Large
GROUP       PUVA         3.078       1.894       5.003    GROUP PUVA


/* Stratified Cox Model Results */
                  Analysis of Maximum Likelihood Estimates
                                 Parameter      Standard
Parameter             DF        Estimate         Error    Chi-Square    Pr > ChiSq
GROUP     PUVA         1         1.10451       0.24783       19.8616        <.0001

                       Hazard      95% Hazard Ratio
Parameter               Ratio      Confidence Limits     Label
GROUP     PUVA          3.018       1.857       4.905    GROUP PUVA
```

Sample Size / Power Calculation for Cox Regression

- Suppose patients are randomized to treatments with $n_1 = rn_2$
- Investigators want to detect the log hazard ratio for treatment of $\beta = \delta$ while adjusting for covariates $\mathbf{X}$
- It can be shown that the required number of total events is

$$d = \frac{(r+1)^2}{r} \frac{(z_{\alpha/2} + z_{\beta})^2}{\delta^2}$$

- Nearly identical formula to that for the log-rank test, only difference being that δ represents the adjusted log HR here
- To determine required number of patients, specify $\varphi = P(\text{patient has event on study})$ and set $n = d/\varphi$

Example

- A clinical trial aims to compare survival rates under standard and new treatment regimens while adjusting for baseline performance status in a Cox model
- Want 90% power to detect an improvement in 5 year survival from 35% under standard to 55% under new treatment
- Equivalently, want to detect
  $\delta = \log HR_{trt} = \log(\log(0.55)/\log(0.35)) = -0.563$
- All patients will be followed for 5 years

Example

- Total number of deaths required is

$$d = \frac{4(1.96 + 1.28)^2}{(-0.563)^2} = 132.475 \approx 133$$

- Under equal allocation, the probability of a patient dying on study can be estimated as
  $\varphi = 0.5 \cdot [1 - S_1(5)] + 0.5 \cdot [1 - S_2(5)] = 0.55$, requiring
  $n = d/\varphi = 242$ patients to be enrolled in total

Example: Sample Size Calculation for Cox Model in SAS

```
proc power;
   coxreg alpha=0.05
      hazardratio = 0.569
      /* SD of treatment indicator */
      stddev = 0.5
      eventprob= 0.55
      power = 0.9
      ntotal = .;
run;
```

The POWER Procedure

Cox Score Test in Proportional Hazards Regression

Fixed Scenario Elements

Method	Hsieh-Lavori normal approximation
Alpha	0.05
Probability of Event	0.55
Test Hazard Ratio	0.569
Test Standard Deviation	0.5
Nominal Power	0.9
Number of Sides	2

Computed N Total

Actual Power	N Total
0.901	241