Chapter 4: Randomization

Instructor: Shu Yang

syang24@ncsu.edu
Department of Statistics
North Carolina State University

2019 Fall Semester

Advantages of randomization

- Eliminates conscious bias
  - physician selection
  - patient self selection
- Balances unconscious bias between treatment groups
  - supportive care
  - patient management
  - patient evaluation
  - unknown factors affecting outcome
- Groups are alike on average
  - allows us to make causal statement for the treatment effect
- Provides a basis for standard methods of statistical analysis such as significance tests

Design-based Inference

Randomization allows us to make *design-based* inference rather than *model-based* inference.

- Suppose we are comparing A to B.
- The **sharp** null H_0: Both treatments A & B would yield exactly the same response on all the patients.
- The design allows us to test the above H_0 without assuming a distribution of data

Design-based Inference

Example

A clinical trial includes 4 patients. Patients 1 & 2 received A, patients 3 & 4 received B.

We would reject H_0 if

$$T = \left(\frac{y_1 + y_2}{2}\right) - \left(\frac{y_3 + y_4}{2}\right)$$

is too large or too small.

- **Question**: How do we calculate the P-value of the above test for given observed T_{obs}?

Design-based Inference

- Under the sharp null, the permutational distribution of T (induced by randomization) can be calculated.

Table: *Permutational distribution under sharp null*

patient	1	2	3	4	
response	y_1	y_2	y_3	y_4	Test statistic T
possible	A	A	B	B	$\left(\frac{y_1+y_2}{2}\right) - \left(\frac{y_3+y_4}{2}\right) = t_1$
treatment	A	B	A	B	$\left(\frac{y_1+y_3}{2}\right) - \left(\frac{y_2+y_4}{2}\right) = t_2$
assignments	A	B	B	A	$\left(\frac{y_1+y_4}{2}\right) - \left(\frac{y_2+y_3}{2}\right) = t_3$
each	B	A	A	B	$\left(\frac{y_2+y_3}{2}\right) - \left(\frac{y_1+y_4}{2}\right) = t_4$
equally	B	A	B	A	$\left(\frac{y_2+y_4}{2}\right) - \left(\frac{y_1+y_3}{2}\right) = t_5$
likely	B	B	A	A	$\left(\frac{y_3+y_4}{2}\right) - \left(\frac{y_1+y_2}{2}\right) = t_6$

- The first one (t_1) is the observed test statistic.
- Under sharp H_0, T can take any of those 6 values with equal prob=1/6.

Design-based Inference

- One-sided P-value (the alternative is "A is better than B"):

$$P[T \geq t_1 \mid \text{sharp } H_0] = \frac{\#\text{ of } t_i \geq t_1}{6}.$$

- Two-sided P-value (the alternative is "A is different than B"):

$$P[|T| \geq |t_1| \mid \text{sharp } H_0] = \frac{\#\text{ of } |t_i| \geq |t_1|}{6}.$$

- When sample size gets large, the distribution looks like normal.
- **Remark**: In the permutational distribution, we treat each individual's response as fixed. Randomness is induced by the treatment assignment mechanism.

Example

suppose A: $y_1 = 8, y_2 = 4$, B: $y_3 = 6, y_4 = 2$. P-values=?

Model-based Inference

- **Statistical model**:

$$Y_1, Y_2 \text{ are iid } N(\mu_1, \sigma^2)$$

$$Y_3, Y_4 \text{ are iid } N(\mu_2, \sigma^2)$$

- Test the null hypothesis

$$H_0 : \mu_1 = \mu_2.$$

- Reject H_0 (in favor of $H_a : \mu_1 > \mu_2$) if the observed

$$T = \frac{\bar{Y}_A - \bar{Y}_B}{s_p(n_A^{-1} + n_B^{-1})^{1/2}}$$

  is too large.
- Under H_0, $T \sim t_{n_A+n_B-2}$.
- The P-value of the above test:

$$\text{P} - \text{value} = P[t_{n_A+n_B-2} \geq T_{\text{obs}}].$$

Model-based Inference

- **Comment**: Distribution-free feature is nice, but the most important aspect of a randomized clinical trial is that it allows us to make **causal** inference. In observational studies such as epidemiological studies, we can only make **associational** statement.

Causal perspective

- Denote treatments A and B as treatments 1 and 2.
- Define potential outcomes $Y^*(1)$ and $Y^*(2)$, where $Y^*(a)$ denotes the outcome had an arbitrary individual in our population been given (possibly contrary to fact) treatment a.
- The individual causal treatment effect of treatment 2 versus treatment 1 is defined as $Y^*(2) - Y^*(1)$.
- The fundamental problem is that only one of the potential outcomes can be observed for a particular individual.
- The average causal treatment effect is

$$\text{ATE} = E\{Y^*(2) - Y^*(1)\} = E\{Y^*(2)\} - E\{Y^*(1)\},$$

where $E\{Y^*(a)\}$ is the mean response in the population had everyone in the population been given treatment a, for $a \in \{1, 2\}$.

Causal perspective

- The observed variable is (A, Y), where A denotes the assigned treatment and Y dentoes the outcome that individual that was assigned treatment A.

Example

In a randomized experiment where n patients (indexed by i) are randomly assigned to receive either treatment 1 or 2 we would observe data which are realizations of (A_i, Y_i), $i = 1, \ldots, n$.

Causal perspective

Assumption (Stable unit treatment value assumption (SUTVA))

The observed outcome Y is the same as the potential outcome for the treatment that was assigned:

$$Y = Y^*(1)I(A = 1) + Y^*(2)I(A = 2)$$

Assumption (Treatment randomization)

The treatment assignment A is statistically independent of the potential outcomes $\{Y^*(1), Y^*(2)\}$.

- This assumption is satisfied in a randomized study because the treatments are assigned through a random mechanism which is certainly independent of an individual's potential outcomes.

Causal perspective

- It follows that $\mu_1 = E(Y \mid A = 1) = E\{Y^*(1)\}$; i.e., the mean response for patients randomized to treatment 1 is equal to $E\{Y^*(1)\}$, because

$$E(Y \mid A = 1) = E\{Y^*(1) \mid A = 1\} = E\{Y^*(1)\},$$

  where the first equality follows from SUTVA and the second equality follows from the independence assumption.
- Similarly, $\mu_2 = E(Y \mid A = 2) = E\{Y^*(2)\}$.
- Therefore, for a randomized study, the treatment difference $\mu_2 - \mu_1$ is the same as the average causal treatment effect.
- In a randomized study, $\bar{Y}_2 - \bar{Y}_1$ is an estimator for the average causal treatment effect.

Disadvantages of Randomization

- Patients or physician may not care to participate in an experiment involving a chance mechanism to decide treatment.
- May interfere with physician patient relationship.
- Part of the resources are expended in the control group; i.e. If we had n patients eligible for a study and had good and reliable historical control data, then it could be more efficient to put all n patients on the new treatment and compare the response rate to the historical controls rather than randomizing the patients into two groups, say, $n/2$ patients on new treatment and $n/2$ on control treatment and then comparing the response rates among these two randomized groups.

How Do We Randomize?

I. Fixed Allocation Randomization

- Consider two treatments A & B.
- If patient population were given A, the average response would be μ_1.
- If patient population were given B, the average response would be μ_2.
- We are interested in estimating $\Delta = \mu_1 - \mu_2$ and make inference on Δ.

I. Fixed Allocation Randomization

- With fixed allocation, the probability that each patient receives treatment A is a constant π, usually $\pi = 0.5$.
- Suppose after treatment allocation, n_1 patients received A and n_2 received B, with $n = n_1 + n_2$ being the total sample size.
- We will estimate Δ using

$$\widehat{\Delta} = \bar{Y}_1 - \bar{Y}_2,$$

$\bar{Y}_1$ is the sample average response of the n_1 patients receiving treatment A, and $\bar{Y}_2$ is the sample average response of the n_2 patients receiving treatment B.

I. Fixed Allocation Randomization

- **When** is $\widehat{\Delta}$ most efficient?
- Assume equal variance, then

$$\text{var}(\widehat{\Delta}) = \text{var}(\bar{Y}_1) + \text{var}(\bar{Y}_2) = \sigma^2 \left(\frac{1}{n_1} + \frac{1}{n_2} \right) \approx \frac{\sigma^2}{n} \left\{ \frac{1}{\pi(1-\pi)} \right\}.$$

- Minimizing the above variance gives $\pi = 0.5$, the minimum variance is $4\sigma^2/n$.
- If $\pi \neq 0.5$, it is less efficient. But the loss is not great. For example, if $\pi = 2/3$, with the same n, the variance of $\text{var}(\widehat{\Delta})$ will be $4.5\sigma^2/n$.
- If we want the same efficiency, we only need to increase sample size by $4.5/4 - 1 = 0.125$. That is, a 12.5% incease in sample size.

I. Fixed Allocation Randomization

Some investigators prefer to put more patients in the new treatment

- better experience on a drug where there is little information
- efficiency loss is slight
- if new treatment is good (as is hoped) more patients will benefit
- might be more cost efficient

Putting more patients in the new treatment has disadvantage too

- might be difficult to justify ethically; It removes equipoise for the participating clinician
- new treatment may be detrimental

Simple randomization

- Each patient has probability π to receive A (hence probability $1 - \pi$ to receive B); usually $\pi = 0.5$.

- For patient i, generate a uniform random variable $U_i \in [0, 1]$

$$\text{If} \begin{cases} U_i \leq \pi \text{ then assign treatment A} \\ U_i > \pi \text{ then assign treatment B.} \end{cases}$$

Advantages of simple randomization

- easy to implement
- virtually impossible for the investigators to guess what the next treatment assignment will be
- the properties of many statistical inferential procedures (tests and estimators) are established under the simple randomization assumption (iid)

Disadvantages of simple randomization

- The major disadvantage is that the number of patients assigned to the different treatments are random. Therefore, the possibility exists of severe treatment imbalance (even with equal allocation probability $\pi = 0.5$)
  - leads to less efficiency:
  - appears awkward and may lead to loss of credibility in the results of the trial

Example

- When $n = 20$,

$$P[\text{imbalance of 12:8 or worse} | \pi = 0.5] \approx 0.5.$$

- When $n = 100$,

$$P[\text{imbalance of 60:40 or worse} | \pi = 0.5] \approx 0.05.$$

Permuted block randomization

- Try to balance A & B.
- Permuted block randomization with a **fixed** block size.

Example (block size=4)

6 possible combinations:

A A B B – per1
A B A B – per2
A B B A – per3
B A A B – per4
B A B A – per5
B B A A – per6

for each block of 4 patients, randomly pick up one combination and assign the treatments to those 4 patients in the sequence specified by the combination.

Ways to Choose a random permutation

1. Order the 6 permutations by $per1 - per6$; generate a uniform random number U_i for the ith block of 4 patients; if $U_i \in [0, 1/6]$, the use $per1$; if $U_i \in [1/6, 2/6]$, then pick up $per2$, etc.

2. For AABB, generate a uniform random number for each letter; re-order the random numbers (ascending or descending). Then the re-ordered letters give a permutation as illustrated in the following table:

Treatment	random number	rank
A	0.069	1
A	0.734	3
B	0.867	4
B	0.312	2

This table gives ABAB.

Potential problem

- If the block size (such as 4) is known, the physician can guess what treatment next patient is going to receive (with certainty for the last treatment). This may cause bias in estimating treatment effect.
- Solution is ...

  Permuted block randomization with **varying block size**: choose several block sizes in advance and pick up a block size with some pre-specified probability; after a block size is chosen, pick up the permutation randomly (with each probability).
- For example, the blocking number may be $2, 4, 6$, chosen at random with, say, each with probability $1/3$. Varying the block sizes at random will make it difficult (not impossible) to break the treatment code.
- The maximum imbalance between A & B: (block size)/2.

Stratified Randomization

- Stratified Randomization (often used with blocking): form strata using prognostic factors; then in each stratum, perform permuted block randomization (with fixed or varying block size).

Example

If age and gender are strong prognostic factors, then we can form following strata:

	Age		
Gender	40-49	50-59	60-69
Male			
Female			

The maximum imbalance between A & B: # of strata $\times$ (block size)/2.

Advantages of Stratified Randomization

- Makes the treatment groups appear similar. This can give more credibility to the results of a study
- Blocked randomization within strata may result in more precise estimates of treatment difference; but one must be careful to conduct the appropriate analysis

Illustration on the effect that blocking within strata has on the precision of estimators

- A prognostic factor with 2 strata:

$$S = \begin{cases} 1 = & \text{strata 1,} \\ 0 = & \text{strata 0.} \end{cases}$$

- Let

$$X = \begin{cases} 1 = & \text{treatment A,} \\ 0 = & \text{treatment B.} \end{cases}$$

- Assume a model for the response Y for the ith patient:

$$Y_i = \mu + \alpha S_i + \beta X_i + \epsilon_i$$

β is the treatment effect, ϵ_i are iid errors with mean 0 and variance σ^2.

- Denote sample means: $\bar{Y}_A$ and $\bar{Y}_B$:

$$\bar{Y}_A = \sum_{X_i=1} Y_i/n_A,$$

$$\bar{Y}_B = \sum_{X_i=0} Y_i/n_B,$$

  where $n_A = \sum_{i=1}^{n} X_i$, number of patients receiving treatment A, $n_B = n - n_A$.

- We will estimate treatment effect β by

$$\widehat{\Delta} = \bar{Y}_A - \bar{Y}_B.$$

Table: *Number of observations falling into the different strata by treatment*

	Treatment		
strata	A	B	total
0	n_{A0}	n_{B0}	n_0
1	n_{A1}	n_{B1}	n_1
total	n_A	n_B	n

- Then

$$\begin{aligned}
\bar{Y}_A &= \sum_{X_i=1} Y_i/n_A \\
&= \sum_{X_i=1} (\mu + \alpha S_i + \beta X_i + \epsilon_i)/n_A \\
&= (n_A\mu + \alpha \sum_{X_i=1} S_i + \beta \sum_{X_i=1} X_i + \sum_{X_i=1} \epsilon_i)/n_A \\
&= (n_A\mu + \alpha n_{A1} + \beta n_A + \sum_{X_i=1} \epsilon_i)/n_A \\
&= \mu + \alpha \frac{n_{A1}}{n_A} + \beta + \bar{\epsilon}_A,
\end{aligned}$$

where $\bar{\epsilon}_A = \sum_{X_i=1} \epsilon_i/n_A$.

- Similarly,

$$\bar{Y}_B = \mu + \alpha \frac{n_{B1}}{n_B} + \bar{\epsilon}_B,$$

  where $\bar{\epsilon}_B = \sum_{X_i=0} \epsilon_i / n_B$.
- Therefore

$$\widehat{\Delta} = \bar{Y}_A - \bar{Y}_B = \beta + \alpha \left(\frac{n_{A1}}{n_A} - \frac{n_{B1}}{n_B} \right) + (\bar{\epsilon}_A - \bar{\epsilon}_B).$$

- Under **stratified blocked randomization**:

$$n_A \approx n_B \approx n/2$$

$$n_{A1} \approx n_{B1} \approx n_1/2$$

$$n_{A0} \approx n_{B0} \approx n_0/2.$$

- So

$$\mathrm{E}(\widehat{\Delta}) = \beta,$$

$$\mathrm{var}(\widehat{\Delta}) = \mathrm{var}(\bar{\epsilon}_A) + \mathrm{var}(\bar{\epsilon}_B) = \sigma^2 \left(\frac{2}{n} + \frac{2}{n} \right) = \frac{4\sigma^2}{n}.$$

- Under **simple randomization**: n_A, n_B, n_{A1} and n_{B1} are all random, and

$$n_{A1}|n_A, n_B \sim b(n_A, \theta), \quad n_{B1}|n_A, n_B \sim b(n_B, \theta),$$

where θ is the probability that a patient is in stratum 1.

- So

$$\begin{aligned}
\mathrm{E}(\widehat{\Delta}) &= \mathrm{E}(\bar{Y}_A - \bar{Y}_B) \\
&= \beta + \alpha \left\{ E\left(\frac{n_{A1}}{n_A}\right) - E\left(\frac{n_{B1}}{n_B}\right) \right\} + E(\bar{\epsilon}_A - \bar{\epsilon}_B) \\
&= \beta
\end{aligned}$$

and

$$\begin{aligned}
\mathrm{var}(\widehat{\Delta}) &= \mathrm{var}(\bar{Y}_A - \bar{Y}_B) \\
&= E\{\mathrm{var}(\bar{Y}_A - \bar{Y}_B|n_A, n_B)\} + \mathrm{var}\{E(\bar{Y}_A - \bar{Y}_B|n_A, n_B)\}.
\end{aligned}$$

- Since

$$
\begin{aligned}
& \mathrm{var}(\bar{Y}_A - \bar{Y}_B | n_A, n_B) \\
= \; & var\left\{ \beta + \alpha \left(\frac{n_{A1}}{n_A} - \frac{n_{B1}}{n_B} \right) + (\bar{\epsilon}_A - \bar{\epsilon}_B) | n_A, n_B \right\} \\
= \; & \alpha^2 \left\{ var\left(\frac{n_{A1}}{n_A} | n_A \right) + var\left(\frac{n_{B1}}{n_B} | n_B \right) \right\} \\
& + \mathrm{var}(\bar{\epsilon}_A | n_A) + \mathrm{var}(\bar{\epsilon}_B | n_B) \\
= \; & \alpha^2 \left\{ \frac{\theta(1-\theta)}{n_A} + \frac{\theta(1-\theta)}{n_B} \right\} + \left(\frac{\sigma^2}{n_A} + \frac{\sigma^2}{n_B} \right) \\
= \; & \{ \sigma^2 + \alpha^2 \theta(1-\theta) \} \left(\frac{1}{n_A} + \frac{1}{n_B} \right).
\end{aligned}
$$

- Therefore

$$\begin{aligned}
\text{var}(\bar{Y}_A - \bar{Y}_B) &= E\{\text{var}(\bar{Y}_A - \bar{Y}_B | n_A, n_B)\} \\
&= \{\sigma^2 + \alpha^2\theta(1-\theta)\}E\left(\frac{1}{n_A} + \frac{1}{n_B}\right) \\
&= \{\sigma^2 + \alpha^2\theta(1-\theta)\}E\left(\frac{1}{n_A} + \frac{1}{n - n_A}\right),
\end{aligned}$$

  where $n_A \sim b(n, 1/2)$.

- It has shown that

$$\frac{1}{n_A} + \frac{1}{n - n_A} \geq \frac{4}{n}.$$

- Hence

$$\text{var}(\widehat{\Delta}) \geq \frac{4}{n}\{\sigma^2 + \alpha^2\theta(1-\theta)\} > \frac{4\sigma^2}{n},$$

  which is the variance $\widehat{\Delta}$ obtained under stratified blocked randomization.

- **Remark:** Suppose we perform stratified blocked randomization, we should take the design into account in data analysis. For example, if we simply use two-sample t-test

$$\frac{\bar{Y}_A - \bar{Y}_B}{s_P \left(\frac{1}{n_A} + \frac{1}{n_B}\right)^{1/2}},$$

where

$$s_P^2 = \left\{ \frac{\sum_{X_i=1}(Y_i - \bar{Y}_A)^2 + \sum_{X_i=0}(Y_i - \bar{Y}_B)^2}{n_A + n_B - 2} \right\}.$$

- It turns out that s_P^2 is an unbiased estimator for $\{\sigma^2 + \alpha^2\theta(1-\theta)\}$ as it should be for simple randomization. However, with stratified randomization, we showed that the variance of $(\bar{Y}_A - \bar{Y}_B)$ is $\frac{4\sigma^2}{n}$.

- Therefore the statistic

$$\frac{\bar{Y}_A - \bar{Y}_B}{s_P \left(\frac{1}{n_A} + \frac{1}{n_B}\right)^{1/2}} \approx \frac{\bar{Y}_A - \bar{Y}_B}{\{\sigma^2 + \alpha^2\theta(1-\theta)\}^{1/2} \left(\frac{2}{n} + \frac{2}{n}\right)^{1/2}},$$

has variance

$$\frac{4\sigma^2/n}{4\{\sigma^2 + \alpha^2\theta(1-\theta)\}/n} = \frac{\sigma^2}{\sigma^2 + \alpha^2\theta(1-\theta)} \leq 1.$$

- Hence the statistic commonly used to test differences in means between two populations

$$\frac{\bar{Y}_A - \bar{Y}_B}{s_P \left(\frac{1}{n_A} + \frac{1}{n_B}\right)^{1/2}},$$

  does not have a t-distribution if used with a stratified design and $\alpha \neq 0$ (i.e. some strata effect). In fact, it has a distribution with smaller variance. Thus, if this test were used in conjunction with a stratified randomized design, then the resulting analysis would be conservative.
- The correct analysis would have considered the strata effect in a two-way analysis of variance ANOVA which would then correctly estimate the variance of the estimator for treatment effect.

- **In general, if we use permuted block randomization within strata in the design, we need to account for this in the analysis.**
- In contrast, if we used simple randomization and the two-sample t-test, we would be making correct inference. Even so, we might still want to account for the effect of strata post-hoc in the analysis to reduce the variance and get more efficient estimators for treatment difference.
- **Disadvantage of blocking within strata**: If we use too many prognostic factors to form strata, we might end up with very few (or even zero) patients in some strata. If each stratum has at most one patient, we are back to simple randomization.

II. Adaptive Randomization Procedures

- Allocation probability depends on the treatment allocation of previous patients

Efron biased coin design

- Choose an integer D and a probability $\phi < 0.5$ (for example, $D = 3$ and $\phi = 0.25$).

- Assign next patient to treatment A with π_A:

$$\begin{aligned} \pi_A &= .5 \quad \text{if} \quad |n_A - n_B| \leq D \\ \pi_A &= \phi \quad \text{if} \quad n_A - n_B > D \\ \pi_A &= 1 - \phi \quad \text{if} \quad n_B - n_A > D \end{aligned}$$

Urn Model (L.J. Wei)

- Start with m balls labeled with A and m balls labeled with B.
- Randomly pick a ball for the first patient and assign the treatment indicated by the ball to that patient.
- If the patient receives A then replace that A ball with a B ball and vise versa.

Minimization Method of Pocock and Simon

- Suppose there are K prognostic factors, each with k_i levels ($i = 1, 2, , , K$).
- At any point in time in the study, let us denote by n_{Aij} the number of patients that are on treatment A for the j-th level of prognostic factor i.
- An analogous definition for n_{Bij}.
- **Note**: If n_A denotes the total number on treatment A, then

$$n_A = \sum_{j=1}^{k_i} n_{Aij}; \text{ for all } i = 1, \ldots, K.$$

Similarly,

$$n_B = \sum_{j=1}^{k_i} n_{Bij}; \text{ for all } i = 1, \ldots, K.$$

- The measure of marginal discrepancy is given by

$$MD = w_0|n_A - n_B| + \sum_{i=1}^{K} w_i(\sum_{j=1}^{k_i} |n_{Aij} - n_{Bij}|).$$

- The weights $w_0, w_1, \ldots, w_K$ are positive numbers which may differ according to the emphasis you want to give to the different prognostic factors. Generally $w_0 = K, w_1 = \ldots = w_K = 1$.
- The next patient that enters the study is assigned either treatment A or treatment B according to whichever makes the subsequent measure of marginal discrepancy smallest. In case of a tie, the next patient is randomized with probability .5 to either treatment.

Example

- Consider two prognostic factors, K=2, the first with two levels, $k_1 = 2$ and the second with three levels $k_2 = 3$.
- Suppose after 50 patients have entered the study, the marginal configuration of counts for treatments A and B, by prognostic factors, looks as follows:

Treatment A				Treatment B			
	PF1				PF1		
PF2	1	2	Total	PF2	1	2	Total
1		*	13	1		*	12
2			9	2			6
3			4	3			6
Total	16	10	26	Total	14	10	24

- If we take the weights to be $w_0 = 2$ and $w_1 = w_2 = 1$, then the measure of marginal discrepancy equals

$$MD = 2|26-24|+1(|16-14|+|10-10|)+1(|13-12|+|9-6|+|4-6|) = 12.$$

- Suppose the next patient entering the study is at the second level of PF1 and the first level of PF2. Which treatment should that patient be randomized to?
- If the patient were randomized to treatment A, then the result would be

Treatment A				Treatment B			
	PF1				PF1		
PF2	1	2	Total	PF2	1	2	Total
1			14	1			12
2			9	2			6
3			4	3			6
Total	16	11	27	Total	14	10	24

and the measure of marginal discrepancy

$$MD = 2|27-24|+1(|16-14|+|11-10|)+1(|14-12|+|9-6|+|4-6|) = 16.$$

- Whereas, if that patient were assigned to treatment B, then

Treatment A				Treatment B			
	PF1				PF1		
PF2	1	2	Total	PF2	1	2	Total
1			13	1			13
2			9	2			6
3			4	3			6
Total	16	10	26	Total	14	11	25

and the measure of marginal discrepancy

$$MD = 2|26-25|+1(|16-14|+|10-11|)+1(|13-13|+|9-6|+|4-6|) = 10.$$

- Therefore, we would assign this patient to treatment B.
- Note that design-based inference is not even possible since the allocation is virtually deterministic.

III. Response Adaptive Randomization

- Allocation probability depends on the **outcome** of the previous patients.

Play-the-Winner Rule (Zelen)

- First patient is randomized to either treatment A or B with equal probability.
- Next patient is assigned the same treatment as the previous one if the previous patient's response was a success; whereas, if the previous patient's response is a failure, then the patient receives the other treatment. The process calls for staying with the winner until a failure occurs and then switching.

Example

	Patient ordering							
Treatment	1	2	3	4	5	6	7	8
A	S	F				S	S	F
B			S	S	F			

Urn Model (L.J. Wei)

- Every time there is a success on treatment A add r A balls into the urn, when there is a failure on treatment A add r B balls. Similarly for treatment B. The next patient is assigned to whichever ball is drawn at random from this urn.
- Response adaptive allocation schemes have the intended purpose of maximizing the number of patients in the trial that receive the superior treatment.

Difficulties with response adaptive allocation schemes

- Information on response may not be available immediately.
- Such strategies may take a greater number of patients to get the desired answer. Even though more patients on the trial may be getting the better treatment, by taking a longer time, this better treatment is deprived from the population at large who may benefit.
- May interfere with the ethical principle of equipoise.
- Results may not be easily interpretable from such a design.

ECMO trial

- Extracorporeal membrane oxygenator was a promising treatment for a neonatal population suffering from respiratory insufficiency. This device oxygenates the blood to compensate for the lung's inability or inefficiency in achieving this task. The mortality rate was very high for this population and due to very promising results of ECMO it was decided to use a play-the-winner rule.
- The first child was randomized to the control group and died. The next 10 children were assigned ECMO and all survived at which point the trial was stopped and ECMO declared a success.
- It turned out that after further investigation, the first child was the sickest of all the children studied. Controversy ensued and the study had to be repeated using a more traditional design.
- Footnote on page 73 of the textbook FFD gives further references.

Mechanics of Randomization

The following formal sequence of events should take place before a patient is randomized into a phase III clinical trial.

- Patient requires treatment
- Patient is eligible for the trial. Inclusion and exclusion criteria should be checked immediately. For a large multi-center trial, this may be done at a central registration office
- Clinician is willing to accept randomization
- Patient consent is obtained. In the US this is a legal requirement
- Patient formally entered into the trial

After a patient and his/her physician agree to participate in the trial then

- Each patient must be formally identified. This can be done by collecting some minimal information; i.e. name, date of birth, hospital number. This information should be kept on a log (perhaps at a central office) and given a trial ID number for future identification. This helps keep track of the patient and it helps guard against investigators not giving the allocated treatment.

- The treatment assignment is obtained from a randomization list. Most often prepared in advance
  1. The randomization list could be transferred to a sequence of sealed envelopes each containing the name of the next treatment on the card. The clinician opens the envelope when a patient has been formerly registered onto the trial
  2. If the trial is double-blind then the pharmacist preparing the drugs needs to be involved. They prepare the sequence of drug packages according to the randomization list.
  3. For a multi-center trial, randomization is carried out by the central office by phone or by computer.
  4. For a double-blind multi-center trial, the randomization may need to be decentralized to each center according to (2). However, central registration is recommended.

Documentation

- A confirmation form needs to be filled out after treatment assignment which contains name, trial number and assigned treatment. If randomization is centralized then this confirmation form gets sent from the central office to the physician. If it is decentralized then it goes from physician to central office.
- An on-study form is then filled out containing all relevant information prior to treatment such as previous therapies, personal characteristics (age, race, gender, etc.), details about clinical conditions and certain laboratory tests (e.g. lung function for respiratory illness)

All of these checks and balances must take place quickly but accurately prior to the patient commencing therapy.