

HAFTA 9

4.4. Rasgele Klinik Denemeler

Amaç: Bir klinik denemenin amacı bir müdahalenin etkilerini değerlendirmektir. Değerlendirme bir karşılaştırma olduğu anlamına gelir. (Yani müdahale yok, en iyi uygun terapi gibi)

Temel prensip: Gruplar sadece her bir grubun aldığı denemelerin farklı olması dışında bütün önemli yönlerde benzer olmak zorundadır. Aksi halde gruplar arasında tedaviye cevap vermede farklılıklar çalışma altındaki tedavilere bağlı olmayabilir, fakat grupların belirli karakteristiklerine dayandırılabilir.

Rasgele olmayan denemeler: Rasgele olmayan denemelerdeki zorluk kontrol grubunun müdahale grubundan farklı olmasıdır. (prognastically) Bu nedenle kontrol grubu ile müdahale grubu arasındaki karşılaştırmalar yanlı olabilir. Ek olarak, bu yanı düzeltme girişimleri ne tasarımı ne de yeterli olmayan analiz (tabakalı analize ayarlanan veya regresyon analizi) gibi diğer faktörlerdeki farklılıklara bağlıdır.

Örnek 4.2. Birleştirilmiş çoklu denemelerle zorluklarını örnekle açıklamak için kalın bağırsağın ileri karsinomu (carcinoma) olan hastalarda hepsinin aynı 5-FU (fluorouracil=fuloruscil) tedavisinde kullanmasındaki 12 çalışmadan sonuçlar tanıtılacaktır.

Tablo 4.8: İleri karsinomun tedavisi için 5-FU'nun hızlı enjeksiyon sonuçları

Grup	Hasta Sayısı	Yanıtların % si
1. Sharps Benefiel	13	85
2. Roohlin et. al.	47	55
3. Cornell et.al.	13	46
4. Field	37	41
5. Weiss Jackson	37	35
6. Hurley	150	31
7. ECOG	48	27
8. Brennan et.al.	183	23
9. Ansfield	141	17
10. Ellison	87	12
11. Knoepp et.al.	11	9
12. Olson and Greene	12	8

Bu tablo yüzde 8 ile 85 arasında aynı tedaviye dayalı bu 12 çalışma için (örnek çapının değişkenliğine dayanan) ampirik yanıt oranlarını içermektedir. Bu çalışmaların aynı kitleden muhtemelen rasgele örneklerden olmadığını önermektedir.

- 5-FU ile karşılaştırma yapılabilmesi için yeni bir tedavinin olduğu söylensin. Bu çalışmalardan elde edilen farklı sonuçlar ışığında, nasıl karşılaştırma yapılabilir?
- Şüphesiz ki, çalışmadan çalışmaya yanıt oranlarının tutarlı olması mümkündür ve bunun doğal örnekleme değişkenliğinin sonucu olacaktır. Eğer çalışmadan çalışmaya değişkenliğin olmadığı kuşkusuz ise yeni bir tedavinin kullanıldığı yeni bir çalışma yürütülebilir ve birlikte alınan cevap oranı ile karşılaştırılır.
- Bu farklılıkların rasgele örneklem dalgalanmasının sonucu veya gerçek çalışmadan çalışmaya farklılıklardan kaynaklanıp kaynaklanmadığı nasıl değerlendirilebilir?

Hiyerarşik (aşamalı) modeller: Vurgulanan farklı cevap oranları ile vurgulanan gruplardan alınan rasgele örneklerden farklı çalışmaların konusu olup olmadığını ele almak amacıyla bir hiyerarşik model kavramı tanıtılır.

- Hiyerarşinin birinci seviyesi:
 - N çalışmanın her biri vurgulanan bir cevap oranına sahip olduğu ve N tane cevap oranı p_1, p_2, \dots, p_N olduğu varsayalım.
 - Bu N çalışmanın varsayımsal bir çalışma kitlesinden örnek alındığı düşünölsün. Böylece çalışma gruplarının bir kitlesinden belirli bir çalışmaya özgü cevap oranlarının bir örneği gibi p_1, p_2, \dots, p_N olarak varsayılır.
 - Özellikle p_1, p_2, \dots, p_N 'in ortalaması μ_p ve varyansı σ_p^2 olan kitleden N çaplı *iid* (bağımsız aynı dağılımlı) bir örnek olduğu varsayılır.
- Hiyerarşinin ikinci seviyesi:
 - n_i = tedavi gören hasta sayısı
 - X_i = tedaviye cevap verenlerin sayısı
 - olmak üzere i çalışmadan elde edilen veri (n_i, X_i) olarak özetlenebilir.
 - n_i ve p_i değerleri verilmişken (bilinmiyor ise)

X_i rasgele değişkeni koşullu Binom dağılımına sahiptir. Yani,

$$X_i | n_i, p_i \sim \text{Binom}(n_i, p_i)$$

Gerçekleşme: Şimdi hiyerarşik model rasgele örneklem varyansları ile gerçek çalışmadan çalışmaya farklılıklar arasında ayırım yapmaya izin verir. Eğer bütün çalışma grupları homojen ise $\sigma_p^2 = 0$ dır. Böylece σ_p^2 parametresinin tahmin edilmesiyle çalışmadan çalışmaya farklılıkların derecesi değerlendirilebilir.

Hatırlatma: Tekrarlanan koşullu beklenen değer

$$E(X) = E[E(X|Y)]$$

ve tekrarlanan koşullu varyans için Adam's kuralı

$$Var(X) = E[Var(X|Y)] + Var[E(X|Y)]$$

dir.

Tahmin: i . çalışmadan örnek oranı $\hat{p}_i = \frac{X_i}{n_i}$ olsun. Beklenen değer ve varyansı

$$E(\hat{p}_i | n_i, p_i) = p_i$$

$$Var(\hat{p}_i | n_i, p_i) = \frac{p_i(1-p_i)}{n_i}$$

Not: Bu problemin kavramlaştırılmasında,

- N çalışma yönlendirilse (çalışmaların bir kitlesinden seçilir.)
- i . çalışmadan n_i çaplı örnek seçilir ve X_i cevap verenlerin sayısı elde edilir.
- n_1, n_2, \dots, n_N örnek çapları bazı dağılımlardan seçilen rasgele değişkenler olduğu varsayalım.

Kavramlaştırma: Bu deneyin sonuçları

$$(n_i, p_i, X_i), \quad i = 1, 2, \dots, N$$

bağımsız aynı dağılımlı (iid) rasgele vektörleri ile özetlenebilir. Şüphesiz ki p_i 'ler gözlemlenemezler. Yine de p_i 'ler hiyerarşik modelin tanımında örtülü (ima edilen) ve önemlidir.

Hatırlatma: Tekrarlanan koşullu beklenen değer ve varyansının kurallarını kullanarak i . çalışma için aşağıdaki sonuçlar elde edilir.

$$E(\hat{p}_i) = E\{E(\hat{p}_i | n_i, p_i)\} = E(p_i) = \mu_p$$

$$\begin{aligned} Var(\hat{p}_i) &= E\{Var(\hat{p}_i | n_i, p_i)\} + Var\{E(\hat{p}_i | n_i, p_i)\} \\ &= E\left\{\frac{p_i(1-p_i)}{n_i}\right\} + Var(p_i) \\ &= E\left\{\frac{p_i(1-p_i)}{n_i}\right\} + \sigma_p^2 \end{aligned}$$

Not: Varyans hesaplanmasında

$$E \left\{ \frac{p_i (1 - p_i)}{n_i} \right\}$$

terimi çalışma içindeki değişkenliği gösterirken σ_p^2 de çalışmadan çalışmaya değişkenliği gösterir.

Tahmin: Bu problemin tanımlanma yolu $\hat{p}_i; i=1,2,\dots,N$ bağımsız aynı dağılımlı rasgele değişkenlerdir. Sonuç olarak, beklenen değer $E(\hat{p}_i) = \mu_p$ 'nin yansız bir tahmin edicisi

$$\bar{p} = \frac{1}{N} \sum_{i=1}^N \hat{p}_i$$

ve varyans $Var(\hat{p}_i)$ 'nın yansız bir tahmin edicisi

$$S_p^2 = \frac{1}{N-1} \sum_{i=1}^N (\hat{p}_i - p_i)^2$$

örnek varyansıdır. $\frac{p_i (1 - p_i)}{n_i}$ için yansız bir tahmin edicisi n_i ve p_i verilmişken $\frac{\hat{p}_i (1 - \hat{p}_i)}{n_i}$ 'nin koşullu beklenen değeri

$$E \left\{ \frac{\hat{p}_i (1 - \hat{p}_i)}{n_i - 1} \middle| n_i, p_i \right\} = \frac{p_i (1 - p_i)}{n_i}$$

ile gösterilebilir. Böylece, $\frac{\hat{p}_i (1 - \hat{p}_i)}{(n_i - 1)}$; $i=1,2,\dots,N$ bağımsız aynı dağılımlı rasgele değişkenler olduğunda

$$E \left\{ \frac{\hat{p}_i (1 - \hat{p}_i)}{n_i - 1} \right\} = E \left\{ E \left[\frac{\hat{p}_i (1 - \hat{p}_i)}{n_i - 1} \middle| n_i, p_i \right] \right\} = E \left\{ \frac{p_i (1 - p_i)}{n_i} \right\}$$

için yansız bir tahmin edici

$$\frac{1}{N} \sum_{i=1}^N \frac{\hat{p}_i (1 - \hat{p}_i)}{n_i - 1}$$

dir.

Özetleme: S_p^2 tahmin edicisinin yansız bir tahmin

$$Var(\hat{p}_i) = E \left\{ \frac{p_i (1 - p_i)}{n_i} \right\} + \sigma_p^2$$

edicisidir. Üstelik $\frac{1}{N} \sum_{i=1}^N \frac{\hat{p}_i(1-\hat{p}_i)}{n_i-1}$ de $E\left\{\frac{p_i(1-p_i)}{n_i}\right\}$ için yansız bir tahmin edicidir. Sonuç olarak σ_p^2 için yansız bir tahmin edici

$$\hat{\sigma}_p^2 = S_p^2 - \frac{1}{N} \sum_{i=1}^N \frac{\hat{p}_i(1-\hat{p}_i)}{n_i-1}$$

dir. Böylece çalışmadan çalışmaya varyans σ_p^2 'nin bir tahmin edicisi olarak elde edilir.

Örnek 4.2 (devam) Tablo 4.8 de verilen 12 çalışmanın sonuçları kullanılarak, ortalama cevap oranı

$$\bar{p} = \frac{1}{12} \sum_{i=1}^{12} \hat{p}_i = 0.3241$$

$$S_p^2 = \frac{1}{12-1} \sum_{i=1}^{12} (\hat{p}_i - \bar{p})^2 = 0.0496$$

ve $\frac{1}{12} \sum_{i=1}^{12} \frac{\hat{p}_i(1-\hat{p}_i)}{n_i-1} = 0.0061$ dir.

Bunun sonucu olarak çalışmadan çalışmaya varyansın tahmini $\hat{\sigma}_p^2 = 0.0496 - 0.0061 = 0.0435$ bulunur. Eğer çalışmadan çalışmaya değişim yoksa, $\hat{\sigma}_p^2 = 0$ olur. Açıkça, burada durum bu değildir. Örnekle açıklama için çalışmadan çalışmaya olan standart sapmanın bir tahmini $\hat{\sigma}_p = \sqrt{0.0435} = 0.2086$ ve bir standart sapma tolerans aralığı

$$\begin{aligned} \bar{p} \pm \hat{\sigma}_p &\Leftrightarrow 0.3241 \pm 0.2086 \\ &\Leftrightarrow (0.1145, 0.5337) \end{aligned}$$

dir. Böylece bu 12 çalışmanın sonuçlarındaki değişkenliğin çok büyük bir kısmıdır.

5. Rasgelelik

Rasgele bir klinik denemede, hastalara tahsis edilen deneme hangi terapinin verildiğini ne hastanın ne de doktorun bilmediği bir şans mekanizması kullanılarak yapılır. Denemedeki her bir hastanın çalışma altında herhangi bir tedaviyi alma şansı aynı olmasıdır.

Rasgeleliğin avantajları:

- Bilinçli yan ortadan kaldırılır. (yani doktor seçimi, hastanın kendi seçimini yapması, vs.)
- Deneme grupları arasında bilinçsiz yanı dengeler (yani, destekleyici bakım, hasta idaresi, hasta değerlendirmesi, sonucu etkileyen bilinmeyen faktörler, vs.)

- Gruplar “ortalama aynı” dır.
- İstatiksel analizin standart yöntemleri için bir temel sağlar. (yani hipotez testleri vs.)

5.1 Tasarım tabanlı sonuç çıkarım:

Sonuç çıkarım: Rasgeleliğin kullanımı model tabanlı sonuç çıkarımı yerine tasarım tabanlı sonuç çıkarımın gerçekleştirilmesine imkân verir. Kısaca rasgele bir denemede olasılık dağılımı rasgeleliğin kendisiyle sonuç çıkarılmasıdır, böylece bir kurumsal (hipoteze bağlı) istatistik modelini belirlemeye gerek yoktur.

Sonuç: İki tedavi denemesinde (yani tedavi A ve tedavi B her bir hastada aynı etkiyi gösterdiği) kesin bir yokluk hipotezi altında bir test istatistiğinin permütasyon dağılımını değerlendirmek mümkündür.

$$H_0 : A \text{ tedavisinin etkisi} = B \text{ tedavisinin etkisi}$$

Örnek 5.1. 4 hastadan ikisine A tedavisi ve diğer ikisine de B tedavisi rasgele atanmış olsun. Bir tedavinin etkisine verilen cevap diğerine verilen cevaptan daha fazla olup olmadığı ilgi alanıdır. Bu tartışmanın hatırlanması tedaviye cevap verme ölçüm olarak alınsın ve Y ile gösterilsin. Örneğin, Y tedavinin başlamasından sonraki bir haftalık kan basıncındaki farklılık olabilir.

- A tedavisini alan hastalarda ortalama cevap ve B tedavisini alan hastalarda ortalama cevap arasındaki farklılığa karşılık gelen bir test istatistiğinin hesaplanmasıyla iki tedavi değerlendirilebilir. Eğer yokluk hipotezi doğru ise bu test istatistiğinin sıfır etrafında bir değer alacağı beklenir. Eğer bu test istatistiği sıfırdan uzak bir değer almışsa, yokluk hipotezine karşı bu bir kanıt olacaktır.
- Olasılık değerleri kanıtın gücünü ölçmek için kullanılacaktır. Yokluk hipotezi altında elde edilen bir test istatistiğinin olasılığı denemede gözlenenlerden daha uç değer aldığı yeterince küçük (< 0.05 diyelim) ise yokluk hipotezinin red edilmesi için bir kanıt olarak kullanılır.

Test istatistiği:

A tedavisini alan iki kişinin tedaviye cevap vermeleri Y_1 ve Y_2 ile, B tedavisini alan iki kişinin de tedaviye cevap vermeleri Y_3 ve Y_4 ile gösterilsin. Test istatistiğinin değeri

$$T = \left(\frac{Y_1 + Y_2}{2} \right) - \left(\frac{Y_3 + Y_4}{2} \right)$$

ile gösterilir. Test istatistiğinin yokluk hipotezinin red edilebilmesi için yeterli kanıtı verip vermeyeceğine nasıl karar verilebilir? Bu hipotez ve rasgelelik altında permütasyon olasılık dağılımı hesaplanabilir. Bu dağılım aşağıda Tablo 5.9 da gösterilebilir.

Tablo 5.9: Örnek 5.1 için keskin bir yokluk hipotezi altında permütasyon olasılık dağılımı

Hasta				Test İstatistiği	Olasılık
1	2	3	4		
A	A	B	B	$T_1 = \frac{1}{2}(Y_1 + Y_2) - \frac{1}{2}(Y_3 + Y_4)$	$\frac{1}{6}$
A	B	A	B	$T_2 = \frac{1}{2}(Y_1 + Y_3) - \frac{1}{2}(Y_2 + Y_4)$	$\frac{1}{6}$
A	B	B	A	$T_3 = \frac{1}{2}(Y_1 + Y_4) - \frac{1}{2}(Y_2 + Y_3)$	$\frac{1}{6}$
B	A	A	B	$T_4 = \frac{1}{2}(Y_2 + Y_3) - \frac{1}{2}(Y_1 + Y_4)$	$\frac{1}{6}$
B	A	B	A	$T_5 = \frac{1}{2}(Y_2 + Y_4) - \frac{1}{2}(Y_1 + Y_3)$	$\frac{1}{6}$
B	B	A	A	$T_6 = \frac{1}{2}(Y_3 + Y_4) - \frac{1}{2}(Y_1 + Y_2)$	$\frac{1}{6}$

Not: Yokluk hipotezi altında test istatistiğinin T_1, T_2, \dots, T_6 değerlerinin her birini $1/6$ olasılıkla alabilir. Gözlenen gerçek değer bu durumda T_1 uyarlanan olasılık dağılımı ile karşılaştırılması sonucu yeterli derecede büyük olup olmadığı açıklanabilir. Bu

$$P(T \geq T_1 | \text{Yokluk hipotezi})$$

olasılığının hesaplanmasıdır. Bu örnekte T_1 'e eşit ve büyük olan T_i 'lerin sayısının altıya bölünmesidir. Bu tek yönlü olasılık değeridir. Yani

$$H_0 : A \text{ ve B tedavisi aynı etkiye sahiptir.}$$

$$H_1 : A \text{ tedavisi B tedavisinden daha iyidir.}$$

hipotezlerinin test edilmesinin ölçüsüdür. Tek yönlü hipotez yerine iki yönlü alternatif hipotez (yani A ve B tedavilerinin etkileri farklıdır) ise $|T|$ değeri büyük değerleri veya eşiti

$$P(|T| \geq T_1 | \text{Yokluk hipotezi})$$

olasılık değerinin küçük değerleri H_0 hipotezine karşı kanıt olarak kullanılabilir. Çünkü T 'nin gerçek permütasyon dağılımının sıfır etrafında simetrik olması gerekli değildir. Bunun anlamı

iki yönlü test için olasılık deęerleri tek yönlü testin olasılık deęerlerinin iki katıdır. (İkinci koşuluyla, şüphesiz ki $1/2$ 'ye eşit ya da daha az olmasıdır.)

Açıklama: Örnek 5.1 (4 denekli) çok gerçekçi değildir. Bu örneğin kullanılmasının nedeni permütasyon numaralandırmasını göstermenin kolay olmasındandır. Yine de büyük bir deneme için böyle bir numaralandırma mümkün ve permütasyon olasılık dağılımı aynen hesaplanabilir.