HAFTA 13

7. Phase III denemelerinde örnek çapının belirlenmesi

Önemli: Klinik denemelerde çalışan istatistikçilerin en önemli sorumluluklarından biri bir çalışma için gerekli olan örnek çapının belirlenmesidir. Örnek çapının belirlenmesi için en yaygın işlem en küçük örnek çapının bulunması ki;

- ilgilenilen farklılığın tespitinde yeterli hassaslık istatiksel yöntemle bulunabilir, bu yeterli güce sahip olunmasıdır.
- bu farklılığın tahmini yeterince hassastır.

7.1 Hipotez Testi:

Tedavi farkı: Tedavi farkının bir ölçümüne karşılık gelen bir kitle parametresi çalışmanın birincil son noktası için tanımlanır. Bu fark parametresi Δ olarak tanımlansın. Örneğin,

$\mu_1 = 1.$ tedavi verilen hastalar için kitle ortalaması

$\mu_2 = 2.$tedavi verilen hastalar için kitle ortalaması

olmak üzere iki tedavi arasındaki ortalama cevap

$$\Delta = \mu_1 - \mu_2$$

tedavi farkının ölçüsüdür. Aynı şekilde bir kitlede oranlar ölçülüyorsa tedavi farklılıklarının bir ölçüsü $\Delta = p_1 - p_2$ 'dir. Bir klinik deneme kitle parametresi Δ üzerinde sonuç çıkarım yapmak üzere kurulur.

Hipotez testi:

Yokluk hipotezi : $H_0 :\ \Delta = 0$

İki yönlü alternatif hipotez : $H_1 :\ \Delta \neq 0$

Tek yönlü alternatif hipotez : $H_0 :\ \Delta < 0$ veya $\Delta > 0$

Veri: Yokluk veya alternatif hipotezin doğru olduğuna ilişkin karar vermek için kişilerden rasgele bir örnek alınır.

$Z_i = i.$ kişinin gözlemi (atanan tedavi, cevap verme gibi) olmak üzere n kişiden alınan veri

$Z = (Z_1, Z_2, \ldots, Z_n)$ dir.

Olasılık modelleri: Bir istatistikçi olarak kitle parametrelerinin terimlerinde Z 'nin dağılımını tanımlayan bir olasılık modelini kanıtlamaktır. Diğer parametrelerin olduğu gibi Δ 'nın tahmininin olasılık dağılımını tanımlamak gereklidir. Diğer parametreler ilgilenilen birincil parametre olmayabilir. Bu yüzden bu parametreler genellikle nuisance (sıkıntı) parametreleridir ve θ ile gösterilir.

Basit açıklama: İki örnek problemi

$Y_i =$ sürekli açıklanan değişken (ur küçülmesi, ölüm zamanı gibi)

$A_i =$ atanan tedavi

$$A_i = \begin{cases} 1, & i.\text{ hasta 1. tedaviye atanmışsa} \\ 2, & i.\text{ hasta 2. tedaviye atanmışsa} \end{cases}$$

olarak tanımlanmak üzere $i.$ kişiye ait veri $Z_i = (Y_i, A_i)$ ile gösterilir.

Model tasarımlı sonuç çıkarımı için istatistiksel model

$$Y_i \,\big|\, \{A_i = 1\} \sim N\left(\mu_1,\ \sigma^2\right)$$
$$Y_i \,\big|\, \{A_i = 2\} \sim N\left(\mu_1 + \Delta,\ \sigma^2\right)$$

olsun. Δ ilgilenilen tedavi farklarını gösteren test parametresi olup, test edilmesi gereken hipotez

$$H_0 :\ \Delta = 0$$
$$H_1 :\ \Delta \neq 0$$

dir. $\theta = \left(\mu_1,\ \sigma^2\right)$ nuisance parametrelerinin bir vektörüdür.

7.1.1. Test istatistikleri:

Test istatistiği: Test edilmek istenen hipotez

$$H_0 :\ \Delta = 0$$
$$H_1 :\ \Delta > 0$$

olsun. Bir özet test istatistiğinde birleştirilen veri, değerinin büyüklüğüne dayalı yokluk ve alternatif hipotezler arasında farkı bulmak için kullanılır.

Test istatistiği genellikle

$$T_n \equiv T_n(Z) = T_n(Z_1, Z_2, \ldots, Z_n)$$

ile gösterilir.

- Test istatistiği ile

  a) T_n'in büyük değerleri H_0'a karşı H_1 lehine kanıttır.

  b) T_n'in olasılık dağılımı H_0 ve H_1 yani $\Delta = 0$ arasındaki sınırda değerlendirilebilir.

- $Z=(Z_1,Z_2,\ldots,Z_n)$ gözlemi olmak üzere soruşturmayı yürüttükten sonra ve $Z=z$ gerçekleştirmesini yaptıktan sonra T'nin gerçekleşen değeri hesaplanabilir, yani

$$t_n \equiv T_n(z)=T_n(z_1,z_2,\ldots,z_n)$$

ve H_0 hipotezi altında T'nin dağılımdaki bu değeri H_0'a karşı güçlü kanıtını değerlendirmek için karşılaştırılır.

- Olasılık değerleri H_0 hipotezine karşın kanıtın gücünü değerlendirmek amacıyla kullanılır.

Not: $H_0:\ \Delta\le 0$
$H_1:\ \Delta>0$

test edilmek istenirse pratikte en çok kullanılan test istatistiği

$$P_\Delta(T_n>x)\equiv P(T_n>x|\Delta)$$

olup, bütün x'ler için Δ'nın artan bir fonksiyonudur. Böylece

$\Delta=0$ olduğunda $P_\Delta(T_n>t_n)<\alpha$ ise bütün $\Delta\le 0$ $P_\Delta(T_n>t_n)<\alpha$ dır. Aynı şekilde,

$$\begin{matrix} H_0:\ \Delta\ge 0 \\ H_1:\ \Delta<0 \end{matrix} \quad \text{veya} \quad \begin{matrix} H_0:\ \Delta=0 \\ H_1:\ \Delta\ne 0 \end{matrix}$$

hipotezleri de alınabilir.

Not: Bir çok test istatistiği T_n, $\Delta=0$ olduğunda nuisance parametresi ne değer alırsa alsın dağılımı standart normal dağılıma sahip olduğunda hesaplanır. Bu durumda, T_n'in dağılımı

$$T_n \overset{\Delta=0}{\sim} N(0,1) \text{ veya } T_n \overset{\Delta=0}{\sim} AN(0,1)$$

olarak yazılır. Bir test istatistiği yaklaşık normal dağılıma sahiptir. Örnek çapının ∞'a doğru büyümesi varsayımı altında asimptotik teoriyle bir test istatistiğinin dağılımının normal dağılıma yakınsamasıdır.

İki örnek problemi: (devam)

İki tedaviye verilen ortalama cevap arasındaki karşılaştırma durumu

$$Y_i\big|\{A_i=1\}\sim N(\mu_1,\ \sigma^2)$$
$$Y_i\big|\{A_i=2\}\sim N(\mu_1+\Delta,\ \sigma^2)$$

olmak üzere

$$\begin{aligned} H_0 &: \Delta \le 0 \\ H_1 &: \Delta > 0 \end{aligned} \quad (\mu_1 = \mu_1 + \Delta \;\Rightarrow\; \Delta = 0)$$

hipotezinin test istatistiği

$$T_n = \frac{\bar{Y}_{1+} - \bar{Y}_{2+}}{S_p\sqrt{\frac{1}{n_1} + \frac{1}{n_2}}}$$

dir. Burada $\bar{Y}_{i+} = i.$ örneğin cevap ortalaması ve pooled (toplanmış) varyans tahmini

$$S_p^2 = \frac{(n_1 - 1)S_1^2 + (n_2 - 1)S_2^2}{n_1 + n_2 - 2}$$

dir. $\Delta = 0$ olduğunda, T_n test istatistiği $(n_1 + n_2 - 2)$ serbestlik dereceli $t-$dağılımına sahiptir. Bununla birlikte, eğer $n = n_1 + n_2 - 2$ büyükse (bu genellikle Phase III denemelerde olduğu gibi) T_n'in dağılımı bir $N(0,1)$ dağılımına yaklaşır.

Red bölgesi: Genelde, bir hipotez testi için red bölgesi yokluk hipotezinin red edilmesine yol açan örnek uzayındaki veri noktalarının oluşturduğu bir kümedir. Amaç:

$$\begin{aligned} H_0 &: \Delta \le 0 \\ H_1 &: \Delta > 0 \end{aligned}$$

hipotezinin test edilmesi ve buna ilişkin test istatistiği $T_n \overset{\Delta=0}{\sim} N(0,1)$ olsun.

H_0 hipotezinin red edilmesi kuralı:

$$P_{\Delta=0}(T_n > t_n) \le \alpha \;\Leftrightarrow\; T_n > z_\alpha \text{ dır.}$$

$z_\alpha =$ standart normal dağılımın üsten $\alpha.$ yüzdelik değeri (tablo veya kritik değer)

Küme $\left\{ \underset{\sim}{Z} = (Z_1, Z_2, \ldots, Z_n) : \; T_n\left(\underset{\sim}{Z}\right) \ge z_\alpha \right\}$

α test düzeyinde red bölgesini gösterir.

Not: Benzer ifadeler diğer testler için geçerlidir. Örneğin,

$$\begin{aligned} H_0 &: \Delta = 0 \\ H_1 &: \Delta \ne 0 \quad (\text{iki yönlü alternatif}) \end{aligned}$$

hipotezinin test istatistiği

$$T_n \overset{\Delta=0}{\sim} N(0,1)$$

dir. Bu iki yönlü test için olasılık değeri

$$P_{\Delta=0}(|T_n| \ge t_n) = P_{\Delta=0}(T_n \ge t_n) + P_{\Delta=0}(T_n \le -t_n)$$

dir ve $z_{\frac{\alpha}{2}} =$ standart normal dağılımın üsten $\frac{\alpha}{2}$ yüzdelik değeri (tablo veya kritik değer) olmak üzere test için red bölgesi

$$\left\{ \underset{\sim}{Z} = (Z_1, Z_2, \ldots, Z_n) : \ \left| T_n\left(\underset{\sim}{Z}\right) \right| \geq z_{\frac{\alpha}{2}} \right\}$$

kümesi ile verilir.

7.1.2 Güç hesabı:

Klinik olarak önemli bir fark algılanması önemli sayılan kitle parametresi Δ'nın minimum değerine karşılık gelir. Klinik olarak önemli fark Δ_A ile gösterilsin. Örneğin, yeni bir tedavi dikkate alınarak, standart bir tedavi ile karşılaştırıldığında cevap oranında 0.15 lik bir artışın tespit edilmesi klinik olarak önemli olduğu düşünülsün. Bu durumda $\Delta_A = 0.15$ olur. Bir hipotez testi yaparken kararın hassaslığı gerçekte H_1 doğru iken H_0 hipotezinin red edilme olasılığı ile hesaplanır. Klinik olarak önemli fark varsa bu testin gücü olarak alınır. Şüphesiz ki $\Delta \geq \Delta_A$ olduğunda güç büyük olacaktır. Testin gücü olarak genellikle 0.80, 0.90 veya 0.95 kullanılır.

Hipotez testine bağlantı:

Tek yönlü bir hipotez testi düşünülsün

$$H_0 : \ \Delta \leq 0$$
$$H_1 : \ \Delta > 0$$

Klinik olarak önemli fark Δ_A tanımlanmasıyla aslında parametre uzayının bölgesi $(0, \ \Delta_A)$ bir farksızlık bölgesidir.

- Gerçekte $\Delta \leq 0$ ise küçük olasılıkla $\leq \alpha$ ile H_0 hipotezi red edilir. Şüphesiz ki α düzeyinde test ediliyorsa bu garantidir.
- Öte yandan $\Delta_A > 0$ olduğunda $\Delta > \Delta_A$ ise klinik olarak önemli fark vardır ki büyük olasılıkla H_0 hipotezi red edilir. Yani bu olasılık güçten büyük veya eşittir.
- Gerçekte $0 < \Delta < \Delta_A$ ise H_0 hipotezi red edilir veya red edilemez kararı verilir. Herhangi bir kararla memnun olunur.

Seviye ve güç:

Amaç: $H_0 : \ \Delta \leq 0$

$H_1 : \ \Delta > 0$

hipotezini test etmektir. Bunun için test istatistiği $T_n \overset{\Delta=0}{\sim} N(0,1)$ dir. Tek yönlü hipotezinin red bölge düzeyi

$$P_{\Delta=0}\left(T_n > z_\alpha\right)$$

$z_\alpha = N(0,1)$ dağılımının sağ uçtan α 'lık düzeyinin kritik değeri

Sıklıkla $\Delta = \Delta_A$ olduğunda T_n bir normal dağılıma sahiptir. Yani H_1 hipotezi doğru ise

$$T_n \overset{H_1}{\sim} N\left\{\delta\left(n, \Delta_A, \theta\right);\ \sigma_*^2\left(\Delta_A, \theta\right)\right\}$$

dır. Başka bir ifade ile $\Delta = \Delta_A$ olduğunda T_n ortalaması $\delta\left(n, \Delta_A, \theta\right)$ ve varyansı $\sigma_*^2\left(\Delta_A, \theta\right)$ olan bir normal dağılıma sahiptir. $\Delta = \Delta_A$ olduğunda testin gücü

$$P_{\Delta=\Delta_A}\left(T_n \geq z_\alpha\right)$$

olup,

$z_\alpha = N\left\{\delta\left(n, \Delta_A, \theta\right);\ \sigma_*^2\left(\Delta_A, \theta\right)\right\}$ dağılımının altında α anlamlılık düzeyini belirleyen kritik değer

Notasyon: H_1 hipotezi altında $\Delta = \Delta_A$ olduğunda $\delta\left(n, \Delta_A, \theta\right) = T_n$ 'nin dağılımının ortalaması, örnek çapı n, klinik olarak önemli fark Δ_A ve nuisance parametresi θ 'ya bağlıdır. Benzer şekilde $\sigma_*^2\left(\Delta_A, \theta\right) = T_n$ 'nin dağılımının varyansı Δ_A ve θ 'ya bağlıdır.

Önemli:

- H_0 hipotezinin aksine, H_1 hipotezinin altında T_n 'nin dağılımı nuisance (sıkıntı) parametresine bağlıdır. Böylece, tasarım aşamasında istenen bir gücü veren gerekli örnek çapını belirlemek için klinik olarak önemli fark Δ_A 'nın yanısıra nuisance parametresinin makul bir değerinin olması gereklidir. Bu θ parametresi için daha önce yapılan çalışmalardan bir tahmin elde edilebilir.
- H_1 hipotezi altında $\sigma_*^2\left(\Delta_A, \theta\right)$ varyansı 1'e eşittir. (veya yaklaşık olarak 1 dir) Bu durumda H_1 altında $\delta\left(n, \Delta_A, \theta\right)$ ortalaması merkezi olmama parametresi olarak adlandırılır.

İki örnek problemi:

İki tedavi arasındaki ortalama cevapların karşılaştırılması durumu için

$$Y_i \left| \left\{ A_i = 1 \right\} \sim N\left(\mu_1,\ \sigma_Y^2 \right) \right.$$
$$Y_i \left| \left\{ A_i = 2 \right\} \sim N\left(\mu_1 + \Delta,\ \sigma_Y^2 \right) \right.$$

Test istatistiği

$$T_n = \frac{\overline{Y}_{1+} - \overline{Y}_{2+}}{S_p \sqrt{\frac{1}{n_1} + \frac{1}{n_2}}} \approx \frac{\overline{Y}_{1+} - \overline{Y}_{2+}}{\sigma_Y \sqrt{\frac{1}{n_1} + \frac{1}{n_2}}}$$

$\Delta = 0$ olması yani $\mu_2 = \mu_1 + \Delta$ durumunda $T_n \sim AN(0,1)$, alternatif hipotez $H_1:\ \Delta = \Delta_A$ altında T_n'nin dağılımı yaklaşık normal dağılıma sahiptir. Ortalaması

$$E_{H_1}(T_n) \approx E_{H_1}\left\{ \frac{\overline{Y}_{1+} - \overline{Y}_{2+}}{\sigma_Y \sqrt{\frac{1}{n_1} + \frac{1}{n_2}}} \right\} = \frac{\mu_1 - \mu_2}{\sigma_Y \sqrt{\frac{1}{n_1} + \frac{1}{n_2}}} = \frac{\Delta_A}{\sigma_Y \sqrt{\frac{1}{n_1} + \frac{1}{n_2}}}$$

ve varyansı

$$Var_{H_1}(T_n) \approx Var_{H_1}\left\{ \frac{\overline{Y}_{1+} - \overline{Y}_{2+}}{\sigma_Y \sqrt{\frac{1}{n_1} + \frac{1}{n_2}}} \right\} = \frac{Var_{H_1}(\overline{Y}_{1+}) + Var_{H_1}(\overline{Y}_{2+})}{\sigma_Y^2 \left(\frac{1}{n_1} + \frac{1}{n_2} \right)} = \frac{\frac{\sigma_Y^2}{n_1} + \frac{\sigma_Y^2}{n_2}}{\sigma_Y^2 \left(\frac{1}{n_1} + \frac{1}{n_2} \right)} = \frac{\sigma_Y^2 \left(\frac{1}{n_1} + \frac{1}{n_2} \right)}{\sigma_Y^2 \left(\frac{1}{n_1} + \frac{1}{n_2} \right)} = 1$$

Böylece, $H_1:\ \Delta = \Delta_A$ doğru olduğunda

$$T_n \overset{H_1}{\sim} AN\left\{ \frac{\Delta_A}{\sigma_Y \sqrt{\frac{1}{n_1} + \frac{1}{n_2}}},\ 1 \right\}$$

dir. Merkezi olmama parametresi ortalamaya eşit olduğundan

$$\delta(n, \Delta_A, \theta) = \frac{\Delta_A}{\sigma_Y \sqrt{\frac{1}{n_1} + \frac{1}{n_2}}} \text{ ve } \sigma_*^2(\Delta_A, \theta) \approx 1 \text{ dir.}$$

Açıklama: Gerçekte, $H_1:\ \mu_1-\mu_2=\Delta=\Delta_A$ altında test istatistiği

$$T_n \ \sim\ t_{n_1+n_2-2;\delta}$$

serbestlik derecesi $\left(n_1+n_2-2\right)$ ve merkezi olmama parametresi $\delta=\delta\left(n,\Delta_A,\theta\right)$ olan merkezi olmayan t dağılımına sahiptir.

Ancak $n\rightarrow\infty$ iken $t_{n_1+n_2-2;\delta}\ \approx\ Normal$ 'dir.