Hipotez Testi

IST 108 Olasılık ve İstatistik

Bahar 2016

Yrd. Doç. Dr. Ferhat Dikbıyık

Hipotez Testi

- Bir popülasyondan seçilmiş rastgele bir örneği kullanarak popülasyonun geneli ile ilgili bir hipotezin test edilmesi.
- İstatistiki hipotez, popülasyon dağılımına ait bir parametreler kümesi ile alakalı ifadelerdir.
- Örneğin, bir inşaat firması yüklü miktarda kablo alımı yapacak ve kablo üreticisi ürettikleri kabloların ortalama kırılma kuvvetlerinin en az 7.000psi olduğunu iddia ediyor.
- Bu iddiayı (hipotezi) test etmek için, inşaat firması 10 tane rastgele kablo örneği alabilir ve bu örnek üzerinde yapacakları testlere göre üreticinin hipotezini kabul edip etmeyeceklerini karar verebilirler.

Hipotez Testi

- Ana problem, bu popülasyondan seçilmiş örneğin belirtilen hipotez ile tutarlı olup olmadığını tespit edecek prosedürleri bulmaktır.
- Örneğin, Normal olarak dağıtılmış bir popülasyonun bilinmeyen bir beklentisi (θ) ve bilinen bir varyansı (1) olsun.
- “θ, 1’den küçüktür” ifadesi, rastgele örneği inceleyerek test edebileceğimiz bir istatistiki hipotezdir. Eğer rastgele örnek hipotez ile uyumlu ise hipotez “kabul edilebilir”, aksi halde “reddedilir”.
- Bir hipotezi kabul ederek, *hipotezin doğruluğunu kabul etmiyoruz, ama hipotezin elimizdeki sonuçlar ile tutarlı olduğunu söylüyoruz*.
- Normal (θ, 1) popülasyondan seçilmiş 10 boyutunda bir rastgele örnek için, eğer bu örneğin ortalaması;
  - 1,25 ise hipotezin faydasına bir kanıt olarak görünmese de hipotez kabul edilebilir.
  - 3 ise, $\theta < 1$ olduğunda dahi seçilen 10 adet rastgele değerin ortalaması 3 olabilir, fakat bu çok küçük bir ihtimaldir. O yüzden elimizdeki sonuç hipotez ile tutarsızdır.

Hipotez çeşitleri

- θ‘nın bilinmediği F_θ dağılımına sahip bir popülasyonda θ ile ilgili spesifik bir hipotezi test etmek istiyoruz. H_0 (*sıfır hipotez* diye adlandırılır) bu testi simgelesin. Aşağıda iki örnek sıfır hipotez verilmiş:
  - $H_0 : \theta = 1$
  - $H_0 : \theta \leq 1$
- Eğer birinci hipotez doğruysa, bu popülasyon dağılımını tümüyle tanımlar. Bu tür testlere *basit hipotez* denir.
- Eğer hipotezin doğruluğu, popülasyon dağılımını tümüyle tanımlamıyorsa (örneğin ikinci hipotez gibi) bu tür hipotezlere *bileşik hipotez* denir.

Kritik Bölge

- Bir sıfır hipotezini test etmek için *n* boyutlu rastgele seçilmiş bir popülasyon örneği, X_1, X_2, …, X_n gözlemlenir ve bu *n* adet değere göre hipotezi kabul edip etmeyeceğimize karar veririz.
- n-boyutlu uzayda bir *C* bölgesi tanımlanabilir, öyle ki eğer rastgele örnek X_1, X_2, …, X_n bu bölgede ise hipotez reddedilecektir. Bu bölgeye ***kritik bölge*** denir.
- Diğer bir deyişle, kritik bölge *C* tarafından belirlenen istatistiki test,
  - H_0'ı kabul eder, eğer $(X_1, X_2, \ldots, X_n)$ *C*'de değil ise
  - H_0'ı reddeder, eğer $(X_1, X_2, \ldots, X_n)$ *C*'de ise

Kritik Bölge

- Örneğin, varyansı 1 olan bir normal dağılımlı popülasyonun beklentisinin 1 olması hipotezi için yaygın test şu şekildedir.

$$C = \left\{ (X_1, \ldots\ldots, X_n) : \left| \frac{\sum_{i=1}^{n} X_i}{n} - 1 \right| > \frac{1{,}96}{\sqrt{n}} \right\}$$

- Bu test $H_0 : \theta = 1$ hipotezini, örnekleme ortalamasının bir eksiği 1,96'nın örnekleme boyutunun kareköküne bölümünden fazla olması halinde reddeder. Ayrıntılar sonraki slaytlarda verilecek.

Hata tipleri

- Hata tipi I: H_0 doğru olduğunda H_0'ı reddetmek
- Hata tipi II: H_0 yanlış olduğunda H_0'ı kabul etmek

Önem seviyeleri (anlam düzeyi)

- Amacımız hipotezin doğru olup olmadığını bulmak değil, rastgele örnek sonuçlarına göre hipotezin tutarlı olup olmadığını tespit etmektir.
- Bu durumda elde ettiğimiz sonuçların gerçekleşmesi ihtimali hipotez doğru olduğunda çok küçük ise, hipotezi reddetmek mantıklıdır.
- Bunu yapmanın bir yolu, bir α değeri belirlemek ve hipotez doğru olduğunda reddedilme (hata tipi I) ihtimalinin α'dan hiçbir zaman büyük olmadığını test etmektir.
- α değeri bir testin *önem seviyesi* olarak adlandırılır, önceden belirlenir ve tipik olarak şu değerleri alır: 0,1; 0,05; 0,005.

Önem seviyeleri (anlam düzeyi)

- Popülasyonun bilinmeyen bir parametresi θ ile ilgili bir hipotez şöyle ifade edilsin.
  - $H_0 : \theta \in w$
- H_0 için α önem seviyesinde bir test geliştirmenin en yaygın yolu bir nokta değerlendirici (point estimator), $d(\mathbf{X})$ belirlemektir. Hipotez, eğer $d(\mathbf{X})$ w bölgesinden "çok uzakta" ise reddedilecektir.
- Fakat, H_0'ı reddedebilmek için $d(\mathbf{X})$ "ne kadar uzakta" olmalı sorusunu cevaplamak gerekir ve bunun için H_0 doğru iken $d(\mathbf{X})$'in dağılıma bakarız. Bu gereken önem seviyesi için kritik bölgeyi belirlememize yardımcı olur.

Normal bir popülasyonun beklentisi ile ilgili testler

- Varyansın bilindiği durum
- Varyansın bilinmediği durum: T-testi (Önümüzdeki hafta)

Varyansın bilindiği durum

- Beklentisi (μ) bilinmeyen ve varyası (σ^2) bilinen bir normal popülasyondan *n* boyutunda bir örnek alalım; $X_1, X_2, \ldots, X_n$.

- Aşağıdaki sıfır hipotezi (H_0), alternatif hipotezine (H_1) karşı test edelim. Burada μ_0 belirli bir sabit sayıdır.
  - $H_0 : \mu = \mu_0$
  - $H_1 : \mu \neq \mu_0$

Varyansın bilindiği durum: Kritik bölgenin bulunması

- Örnekleme ortalaması, beklentinin (μ) doğal bir nokta değerlendiricisi olduğuna göre, eğer μ_0, ortalamadan çok uzakta değilse H_0 kabul edilebilir. Bu durumda kritik bölge iyi seçilmiş bir *c* değeri için;

$$C = \{(X_1, \ldots\ldots, X_n) : |\overline{X} - \mu_0| > c\}$$

- Önem seviyesi *α* olduğuna göre, kritik değer *c*'yi bulmak için μ = μ_0 olduğunda Hata Tipi I ihtimalini *α*'ya eşitleriz.

$$P_{\mu_0}\{|\overline{X} - \mu_0| > c\} = \alpha$$

Varyansın bilindiği durum : Kritik bölgenin bulunması

- μ = μ_0 olduğunda, örnekleme ortalaması, beklentisi μ_0 ve varyansı σ^2/n olan bir normal dağılıma sahip olacaktır. Ortalamayı bir standart normal dağılıma sahip rastgele değişkene (Z) dönüştürürsek;

$$Z \equiv \frac{\overline{X} - \mu_0}{\sigma / \sqrt{n}}$$

$$P\left\{\left|\overline{X} - \mu_0\right| > c\right\} = \alpha \Rightarrow P\left\{|Z| > \frac{c\sqrt{n}}{\sigma}\right\} = \alpha$$

$$2P\left\{Z > \frac{c\sqrt{n}}{\sigma}\right\} = \alpha \Rightarrow P\left\{Z > \frac{c\sqrt{n}}{\sigma}\right\} = \alpha / 2$$

Standart Normal Dağılımla ilgili bir not

- Tam bu noktada standart normal dağılım ile ilgili kullanabileceğimiz önemli bir özellikten bahsedelim.
- $a \in (0, 1)$ için, bir z_a değeri olsun öyle ki;

$$P\{Z > z_a\} = 1 - \Phi(z_a) = a$$

- z_a değeri, her hangi bir a değeri için cdf tablosu kullanılarak bulunabilir. Örneğin

$1 - \Phi(2{,}33) = 0{,}01$ olduğundan $z_{0,01} = 2{,}33$ diyebiliriz.

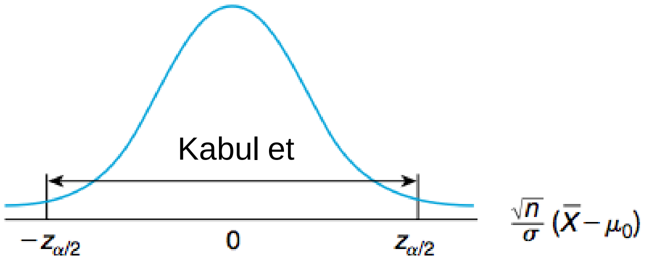
- z_a ile ilgili şu da yazılabilir.

$$P\{Z < z_a\} = 1 - a$$

Varyansın bilindiği durum : Kritik bölgenin bulunması

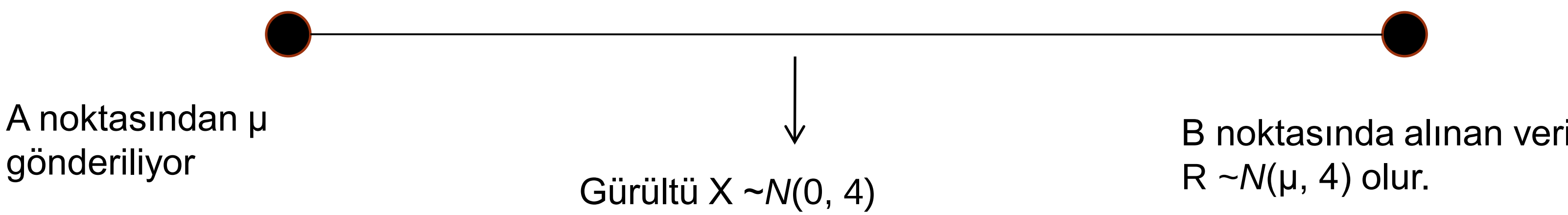
$$P\left\{ Z > \frac{c\sqrt{n}}{\sigma} \right\} = \alpha / 2 \Rightarrow \frac{c\sqrt{n}}{\sigma} = z_{\alpha/2}$$

$$c = \frac{\sigma z_{\alpha/2}}{\sqrt{n}}$$



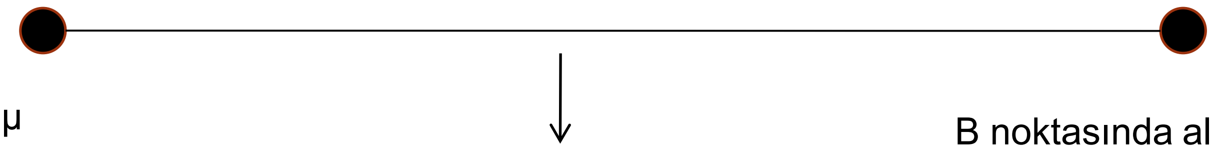
- Yani H_0'ı;
  - reddet et, eğer $\frac{\sqrt{n}}{\sigma}\left|\overline{X} - \mu_0\right| > z_{\alpha/2}$
  - kabul et, eğer $\frac{\sqrt{n}}{\sigma}\left|\overline{X} - \mu_0\right| \leq z_{\alpha/2}$

Örnek 1



B noktasına bugün gönderilecek sinyalin 8 olacağı tahmin ediliyor (H_0 : μ = 8). Eğer aynı sinyal birbirinden bağımsız olarak 5 kere gönderilmişse ve alınan sinyallerin ortalaması 9,5 ise bu hipotezi test edin.

Örnek 1



A noktasından μ gönderiliyor

Gürültü X ~N(0, 4)

B noktasında alınan veri R ~N(μ, 4) olur.

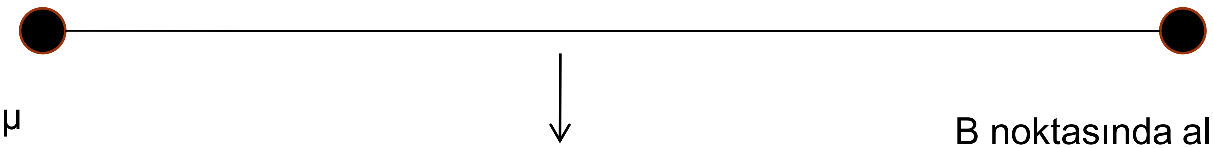
B noktasına bugün gönderilecek sinyalin 8 olacağı tahmin ediliyor (H_0 : μ = 8). Eğer aynı sinyal birbirinden bağımsız olarak 5 kere gönderilmişse ve alınan sinyallerin ortalaması 9,5 ise bu hipotezi test edin.

Önce α = 0,05 için test edelim (yani hipotez doğru iken reddetme ihtimali %5).

$$\frac{\sqrt{n}}{\sigma}\left|\overline{X}-\mu_0\right| = \frac{\sqrt{5}}{2}\left|9,5-8\right| = 1,68 \qquad z_{\alpha/2} = z_{0,025} = 1,96$$

1,68 değeri 1,96 değerinden küçük olduğu için hipotez kabul edilir.

Örnek 1



A noktasından μ gönderiliyor

Gürültü X ~N(0, 4)

B noktasında alınan veri R ~N(μ, 4) olur.

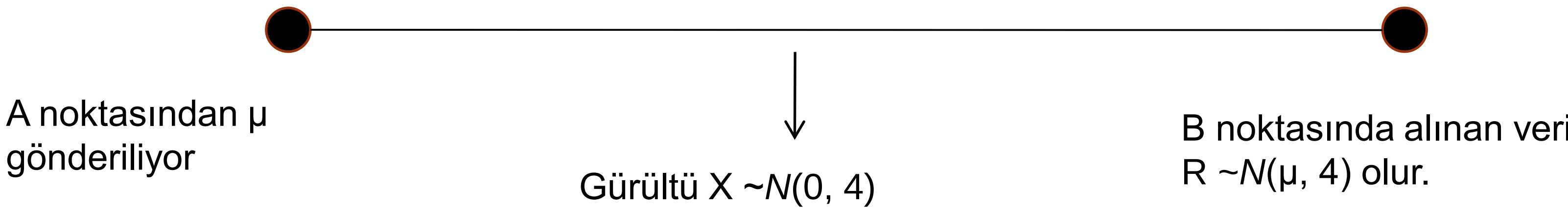
B noktasına bugün gönderilecek sinyalin 8 olacağı tahmin ediliyor (H_0 : μ = 8). Eğer aynı sinyal birbirinden bağımsız olarak 5 kere gönderilmişse ve alınan sinyallerin ortalaması 9,5 ise bu hipotezi test edin.

Şimdi de α = 0,1 için test edelim (yani hipotez doğru iken reddetme ihtimali %10).

$$\frac{\sqrt{n}}{\sigma}\left|\overline{X} - \mu_0\right| = \frac{\sqrt{5}}{2}\left|9{,}5 - 8\right| = 1{,}68 \qquad z_{\alpha/2} = z_{0{,}05} = 1{,}645$$

1,68 değeri 1,645 değerinden büyük olduğu için hipotez reddedilir.

Örnek 1



B noktasına bugün gönderilecek sinyalin 8 olacağı tahmin ediliyor (H_0 : μ = 8). Eğer aynı sinyal birbirinden bağımsız olarak 5 kere gönderilmişse ve alınan sinyallerin ortalaması 9,5 ise bu hipotezi test edin.

Doğru önem seviyesini bulmak önemlidir. Fakat bu, birçok etkene göre değişir. Örneğin, eğer hipotez doğru iken hipotezi reddetmek büyük zarara yol açacaksa hata ihtimalinin düşük olduğu bir önem seviyesi (0,05 veya 0,01 gibi) seçebiliriz.

Varyansın bilindiği durum: Hata Tipi II ihtimali ve OC eğrisi

- Hata Tipi II, $\mu \neq \mu_0$ olduğunda testi kabul etme, ihtimali μ değerine bağlıdır.

$$\beta(\mu) = P_{\mu}\left\{\left|\frac{\overline{X}-\mu_0}{\sigma/\sqrt{n}}\right| \leq z_{\alpha/2}\right\}$$

$$= P_{\mu}\left\{-z_{\alpha/2} \leq \frac{\overline{X}-\mu_0}{\sigma/\sqrt{n}} \leq z_{\alpha/2}\right\}$$

- β(μ) fonksiyonu işletim karakteristiği (operating characteristics - OC) eğrisi olarak adlandırılır ve gerçek beklenti μ iken H_0'ın kabul edilme ihtimalini verir.
- 1-β(μ) *güç fonksiyonu* olarak adlandırılır ve verilen bir μ değeri için, testin gücü, μ gerçek değeri iken reddedilme ihtimalidir.

Varyansın bilindiği durum: Hata Tipi II ihtimali ve OC eğrisi

$$\beta(\mu) = P_\mu\left\{-z_{\alpha/2} \le \frac{\overline{X}-\mu_0}{\sigma/\sqrt{n}} \le z_{\alpha/2}\right\}$$

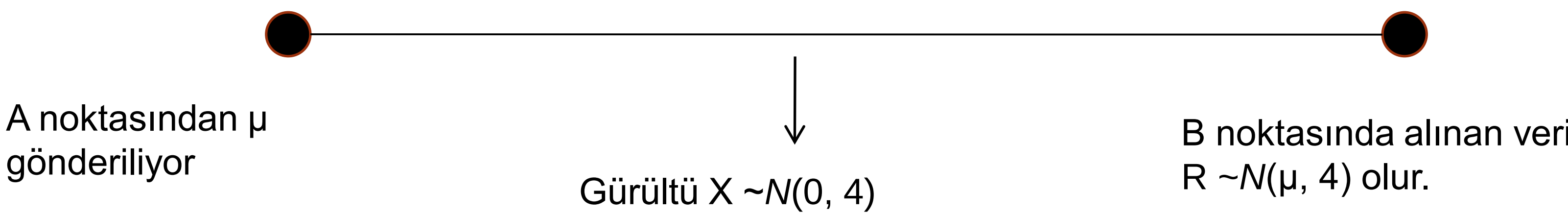
$$= P_\mu\left\{-z_{\alpha/2} - \frac{\mu}{\sigma/\sqrt{n}} \le \frac{\overline{X}-\mu_0-\mu}{\sigma/\sqrt{n}} \le z_{\alpha/2} - \frac{\mu}{\sigma/\sqrt{n}}\right\}$$

$$= P_\mu\left\{-z_{\alpha/2} - \frac{\mu}{\sigma/\sqrt{n}} \le Z - \frac{\mu_0}{\sigma/\sqrt{n}} \le z_{\alpha/2} - \frac{\mu}{\sigma/\sqrt{n}}\right\}$$

$$= P_\mu\left\{\frac{\mu_0-\mu}{\sigma/\sqrt{n}} - z_{\alpha/2} \le Z \le \frac{\mu_0-\mu}{\sigma/\sqrt{n}} + z_{\alpha/2}\right\}$$

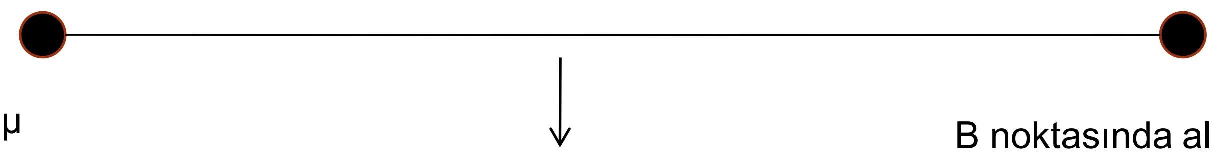
$$= \Phi\left(\frac{\mu_0-\mu}{\sigma/\sqrt{n}} + z_{\alpha/2}\right) - \Phi\left(\frac{\mu_0-\mu}{\sigma/\sqrt{n}} - z_{\alpha/2}\right)$$

Örnek 2



B noktasına bugün gönderilecek sinyalin 8 olacağı tahmin ediliyor (H_0 : μ = 8). Eğer aynı sinyal birbirinden bağımsız olarak 5 kere gönderilmişse ve alınan sinyallerin ortalaması 9,5'dir. Gerçekte gönderilen sinyal 10 ise, Hata Tipi II'yi (yani hipotez yanlışken kabul etme ihtimalini), önem seviyesi 0,05 için hesaplayın.

Örnek 2



A noktasından μ gönderiliyor

Gürültü X ~N(0, 4)

B noktasında alınan veri R ~N(μ, 4) olur.

B noktasına bugün gönderilecek sinyalin 8 olacağı tahmin ediliyor (H_0 : μ = 8). Eğer aynı sinyal birbirinden bağımsız olarak 5 kere gönderilmişse ve alınan sinyallerin ortalaması 9,5'dir. Gerçekte gönderilen sinyal 10 ise, Hata Tipi II'yi (yani hipotez yanlışken kabul etme ihtimalini), önem seviyesi 0,05 için hesaplayın.

$$\beta(\mu) = \Phi\left(\frac{\mu_0 - \mu}{\sigma / \sqrt{n}} + z_{\alpha/2}\right) - \Phi\left(\frac{\mu_0 - \mu}{\sigma / \sqrt{n}} - z_{\alpha/2}\right)$$

$$= \Phi\left(\frac{8 - 10}{2 / \sqrt{5}} + z_{0,025}\right) - \Phi\left(\frac{8 - 10}{2 / \sqrt{5}} - z_{0,025}\right)$$

$$= \Phi\left(-\sqrt{5} + 1{,}96\right) - \Phi\left(-\sqrt{5} - 1{,}96\right) = 0{,}392$$

Varyansın bilindiği durum: Tek-taraflı test

- $\mu = \mu_0$ sıfır hipotezini test ederken, örnekleme ortalaması μ_0'dan uzakta olduğunda reddeden bir test uyguladık.
- Peki $\mu = \mu_0$ olmasının tek karşı alternatifi $\mu > \mu_0$ ise ne olur? Yani $H_0 : \mu = \mu_0$ hipotezine karşı alternatif hipotez $H_1 : \mu > \mu_0$ ise ne olur?
- Bu durumda, örnekleme ortalaması μ_0'dan çok çok büyük olduğunda sıfır hipotezini reddederiz.

$$C = \{(X_1, \ldots, X_n) : \bar{X} - \mu_0 > c\}$$

Varyansın bilindiği durum: Tek-taraflı test – Kritik Bölge

$$C = \{(X_1, \ldots\ldots, X_n) : \overline{X} - \mu_0 > c\}$$

$$P\{\overline{X} - \mu_0 > c\} = \alpha \Rightarrow P\left\{Z > \frac{c\sqrt{n}}{\sigma}\right\} = \alpha$$

$$P\{Z > z_\alpha\} = \alpha \Rightarrow \frac{c\sqrt{n}}{\sigma} = z_\alpha$$

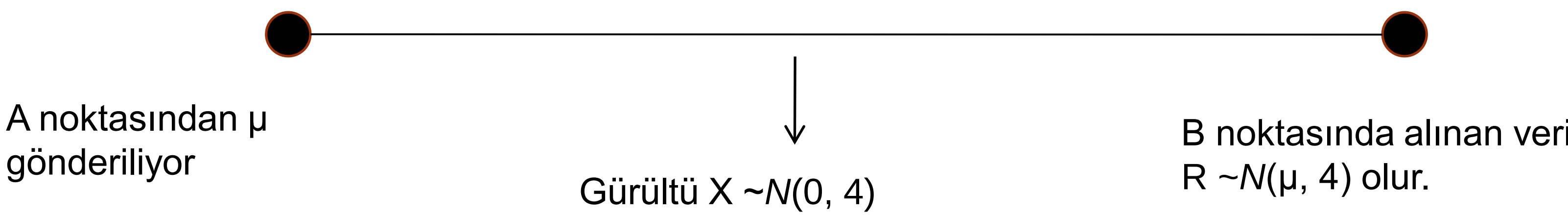
$$c = \frac{\sigma z_\alpha}{\sqrt{n}}$$

Varyansın bilindiği durum: Tek-taraflı test

- Hipotez test problemi:
  - $H_0 : \mu = \mu_0$
  - $H_1 : \mu > \mu_0$

- Çözüm: H_0'ı;
  - reddet et, eğer $\frac{\sqrt{n}}{\sigma}(\overline{X} - \mu_0) > z_\alpha$
  - kabul et, eğer $\frac{\sqrt{n}}{\sigma}(\overline{X} - \mu_0) \leq z_\alpha$

Örnek 3



B noktasına bugün gönderilecek sinyalin 8 olacağı tahmin ediliyor (H_0 : μ = 8). Eğer aynı sinyal birbirinden bağımsız olarak 5 kere gönderilmişse ve alınan sinyallerin ortalaması 9,5'dir. Eğer gönderilen sinyalin en az 8 olduğunu biliyor isek, önem seviyesi 0,05 için hipotezi test edin.

$H_0 : \mu = 8$
$H_1 : \mu > 8$

$$\frac{\sqrt{n}}{\sigma}(\overline{X} - \mu_0) = \frac{\sqrt{5}}{2}(9{,}5 - 8) = 1{,}68 \qquad z_\alpha = z_{0{,}05} = 1{,}645$$

1,68 değeri 1,645 değerinden büyük olduğu için hipotez reddedilir.

Varyansın bilindiği durum: Tek-taraflı test

- Hipotez test problemi:
  - $H_0 : \mu \leq \mu_0$
  - $H_1 : \mu > \mu_0$
- Bu hipotez test problemi de bir önceki ile aynı çözüme sahiptir.
- Çözüm: H_0'ı;
  - reddet et, eğer $\frac{\sqrt{n}}{\sigma}(\overline{X} - \mu_0) > z_\alpha$
  - kabul et, eğer $\frac{\sqrt{n}}{\sigma}(\overline{X} - \mu_0) \leq z_\alpha$

Varyansın bilindiği durum: Tek-taraflı test

- Hipotez test problemi:
  - $H_0 : \mu = \mu_0$ (veya $\mu \geq \mu_0$)
  - $H_1 : \mu < \mu_0$

- Çözüm: H_0'ı;
  - reddet et, eğer $\dfrac{\sqrt{n}}{\sigma}(\overline{X} - \mu_0) < -z_\alpha$
  - kabul et, eğer $\dfrac{\sqrt{n}}{\sigma}(\overline{X} - \mu_0) \geq -z_\alpha$

Örnek 4

- Piyasadaki tüm sigaraların içerdiği ortalama nikotin miktarı sigara başına en az 1,6mg.'dır. Bir sigara üreticisi tütün yapraklarını işlemek için yeni bir metot bulduğunu ve bu sayede sigara başına ortalama nikotin miktarını 1,6mg.'ın altına indirdiğini iddia etmektedir. Bu iddiayı test etmek için bu üreticiden alınan 20 adet sigara analiz edilmiştir. Bir sigaranın nikotin içermesi ile ilgili standart sapmanın 0,8mg. olduğu biliniyorsa ve 20 sigaranın ortalama nikotin miktarı 1,54mg. ise %5'lik önem seviyesi için nasıl bir sonuç çıkarılabilir?

Örnek 4

- Piyasadaki tüm sigaraların içerdiği ortalama nikotin miktarı sigara başına en az 1,6 mg.'dır. Bir sigara üreticisi tütün yapraklarını işlemek için yeni bir bulduğunu ve bu sayede sigara başına ortalama nikotin miktarını 1,6 mg.'ın altına indirdiğini iddia etmektedir. …
- *Üreticinin iddiasını sıfır hipotezi olarak mı yoksa karşı alternatif hipotez olarak mı almalıyız?*
- Sıfır hipotezinin reddi, verinin hipotez ile uyuşmadığının güçlü bir ifadesidir. Fakat hipotezin kabulü durumunda verinin aynı güçlülükle hipotez ile uyuştuğunu söyleyemeyiz.
  - Bu durumda üreticinin iddiasını karşı alternatif hipotez olarak almak mantıklıdır, çünkü eğer sıfır hipotezini reddedersek, güçlü bir şekilde karşı hipotezi kabul etmiş oluruz.
  - H_0 : μ ≥ 1,6
  - H_1 : μ < 1,6

Örnek 4

- Piyasadaki tüm sigaraların içerdiği ortalama nikotin miktarı sigara başına en az 1,6 mg.'dır. Bir sigara üreticisi tütün yapraklarını işlemek için yeni bir bulduğunu ve bu sayede sigara başına ortalama nikotin miktarını 1,6 mg.'ın altına indirdiğini iddia etmektedir. Bu iddiayı test etmek için bu üreticiden alınan 20 adet sigara analiz edilmiştir. Bir sigaranın nikotin içermesi ile ilgili standart sapmanın 0,8 mg. Olduğu biliniyorsa ve 20 sigaranın ortalama nikotin miktarı 1,54mg. ise %5'lik önem seviyesi için nasıl bir sonuç çıkarılabilir?
  - H_0 : μ ≥ 1,6
  - H_1 : μ < 1,6

$$\frac{\sqrt{n}}{\sigma}(\overline{X} - \mu_0) = \frac{\sqrt{20}}{0,8}(1,54 - 1,6) = -0,336 \qquad -z_{\alpha} = -z_{0,05} = -1,645$$

- -0,336 değeri -1,645 değerinden büyük olduğu için sıfır hipotezi kabul edilir, yani karşı hipotezi (üreticinin iddiasını) kabul etmek için yeterli kanıt yoktur.