

199. Tek Örneklem T Testi

Popülasyon ortalaması ile varsayımsal bir değer arasında istatistiksel olarak anlamlı bir farklılık olup olmadığını test etmek için kullanılan parametrik bir testtir. Yani, örnek ortalamasına ilişkin test yapmak için kullanılır.

Daha önce, isledığımız örneklem bölümünde 10.000 kişilik bir ilde 100 tane örnek çekmiştik. Bu çekmiş olduğumuz örneğe **tek örnek** denir. Eğer bu örnek üzerinden ora kitleye ilişkin bir değerlendirme yapmak istiyorsak ki burada popülasyon ortalamasına karşılık geliyor bu. Bu ora kitle ortalamasına ilişkin teorik bir değer ile bir test yapmak istemek bu durumda **tek örneklem t testi** yapmış oluyoruz.

Tek Örneklem T Testi: Hipotezler

$$H_0: \mu = 50$$

$$H_0: \mu \leq 50$$

$$H_0: \mu \geq 50$$

$$H_1: \mu \neq 50$$

$$H_1: \mu > 50$$

$$H_1: \mu < 50$$

H_0 'lar dayanak noktası gibi düşünülmeli.

Tek Örneklem T Testi: Test İstatistiği ***

Tek örneklem T testini yapmak istediğimizde t veya z dağılımlarını kullanabiliyoruz.

Tek örneklem t testinde test istatistiğine karar vermek için yapılacaklar

1. Anakütle standart sapması biliniyorsa z istatistiği kullanılır.
2. Anakütle standart sapması bilinmiyorsa ve $n > 30$ ise z istatistiği kullanılır.
3. Anakütle standart sapması bilinmiyor ve $n < 30$ ise t istatistiği kullanılır.

Fakat veri bilimi kapsamında genelde elimizdeki $n, 30$ 'dan büyük olduğundan dolayı ve örnek sayısı (n) arttıkça t dağılımı normal dağılıma yaklaşacak olduğundan dolayı bu sebeple sürekli t dağılımı düşünülür. Pratikte R ve python programlarının sayesinde zaten kullanılan test istatistiği de genelde t olmuş oluyor.

$$t = \frac{\bar{X} - \mu}{\frac{s}{\sqrt{n}}}$$

$$z = \frac{\bar{X} - \mu}{\frac{\sigma}{\sqrt{n}}}$$

Tek Örneklem T Testi: Varsayım

Hemen hemen bütün hipotez testlerimizde bazı varsayımlarımız olacak ve bütün bu varsayımları test etmemiz gerektiği gerçeği olacak. Bizler bu varsayımları test ederek eğer varsayım sağlanıyorsa ilgili hipotez testini kullanacağız, eğer varsayım sağlanmıyorsa non-parametrik hipotez testlerini ele alacağız. Bütün hipotez testi diyerek ele aldığımız bütün hipotez testleri aslında parametrik hipotez testleri. Fakat varsayılan ön koşullar sağlanmadığında bunları kullanamıyor oluyoruz. Ki veri bilimi gerçek dünyaya uygulandığında genelde non-parametrik durumlar çok oluyor.

Tek örneklem t testinin varsayımı: Normal dağılım olması.

Özetleyecek olursak;

T testi nedir? Elimizde tek bir örneklemin ortalamasına ilişkin test yapma ihtiyacı olduğunda kullanılan testtir. Örneğin bir ülkenin yaş ortalamasıyla ilgili bir hipotez testi kurmak istediğimizde bunu kullanırız. Örneğin, web sitemizde insanların geçirdiği süreyle ilgili bir hipotez testi kurmak istediğimizde 180 saniyedir ya da değildir, 180 saniyeden fazladır, 180 saniyeden azdır şeklinde bir hipotez testi kurmak istediğimizde (elimizde tek bir örnek/tek bir grup/tek bir değişken olduğundan dolayı) bu durumda tek örneklem t testi kullanılır.