BÖLÜM 11

HİPOTEZ TESTLERİ

Önceki bölümde çeşitli parametrelerin örnekleme dağılımlarını ve bunlara dayanarak aralık tahminlerini açıkladık. Bu bölümde açıklayacağımız hipotez testlerinde de aynı örnekleme dağılımlarını kullanacağız.

11.1. HİPOTEZ VE HİPOTEZ TESTİ KAVRAMLARI

Uygulamada anakütle parametreleri ile ilgili bazı varsayımların veya bilgilerin ortaya konduğu durumlarla karşılaşılabilir. Bu varsayımlar veya bilgiler anakütle ortalaması, oranı veya varyansının herhangi bir sayısal değere eşitlik veya eşitsizliğini ifade edebileceği gibi, iki farklı anakütlenin parametrelerinin birbirine eşit veya farklı olduğunun öne sürülmesi şeklinde de karşımıza çıkabilir. Bu durumda ileri sürülen varsayımın veya bilginin doğrulanması gerekecektir.

Bu soruna bir örnek yardımı ile cevap verilebilir. Eğer örnekten tahmin edilen değer ile ileri sürülen değer aynı ise bir sorun yoktur. Buradaki aynı kelimesi tam olarak sayısal eşitliği ifade etmektedir. Sayısal değerler arasında fark olması durumunda ise varsayım yanlıştır demek doğru mudur sorusunu sormak gerekir. Bilindiği gibi örneklemeden doğan bir sapma sözkonusudur. Yani, örnekleme hatası nedeniyle parametrenin tahmin edilen değeri kontrol edilebilir bir miktarda farklı olabilir. Dikkat edilirse burada söz edilen fark, aralık tahmini ile tahmin edilmeye çalışılan farka benzemektedir. Aralık tahmininde fark tahmin edilmeye, yani anakütle parametresinde oluşabilecek sapma tahmin edilmeye çalışılmaktadır. Söz ettiğimiz durumda ise varsayılan bir değer ile örnekten elde edilen bir değer vardır ve bu değerler arasındaki farkın örneklemeden kaynaklanıp kaynaklanmadığının araştırılması sözkonusudur. Diğer bir ifade ile aradaki farkın istatistiksel olarak anlamlı olup olmadığına karar verilmesi sözkonusudur. Bu amaçla hipotez testi olarak adlandırılan işlemler ile değerler arasındaki farkın istatistiksel olarak anlamlılığı araştırılmaktadır.

Hipotez testini, anakütle ile ilgili bilgiler veya varsayımlar ile bunların örnekten tahmin edilen değerleri arasındaki farkın istatistiksel olarak anlamlı olup olmadığına karar vermek için yapılan işlemler olarak tanımlanabilir. Buradaki varsayımlar veya bilgiler, kişisel görüşler olabileceği gibi yapılan başka araştırmalardan elde edilmiş veya başka kaynaklardan alınmış bilgiler de olabilir.

Hipotez testleri, örnekten elde edilecek bilgilerle anakütleye ait bilgilerin karşılaştırılması; örneklerden tahmin edilen iki anakütleye ait bilgilerin karşılaştırılması gibi amaçlarla uygulanacak ve yanlış karar alınmasını önleyici bir kriter sağlayacaktır. Hipotez testlerinin, bu nedenle karar almada önemli yeri vardır. Test sonuçlarının, tahminde olduğu gibi doğruluğunu saptamak güç ise de alınacak kararlarda yardımcı olacağı ve yanlış karar alınmasını önleyeceği unutulmamalıdır.

Bir makinenin ayarının bozulup bozulmadığının, bir fabrikada üretiminin kalitesinde değişme olup olmadığının, yeni geliştirilen bir mamülün kullanılan ile farkı olup olmadığının araştırılması gibi pek çok durumda örnek alınarak anakütlenin bilinen değerleri ile karşılaştırılması gerekebilir. Anakütle parametresinin değeri ile örnekten tahmin edilen değer arasında fark varsa, bunun anakütle parametresinin değeri ile örnekten tahmin edilen değer arasındaki farklılıktan mı kaynaklandığı, yoksa tesadüfi olarak örneklemeden mi kaynaklandığı test ile belirlenebilir. Benzer şekilde, karar almak için iki anakütle parametre tahmininin karşılaştırılması da gerekli olabilir. Örneğin, A ve B gibi iki fabrika aynı malı üretiyorsa, üretimle ilgili herhangi bir özelliği karşılaştırılmak için, her ikisinden de birer örnek alınarak, örnekten tahmin edilen değerler arasında fark varsa, bu farkın A ve B fabrikalarının üretim yapılarındaki özelliklerden mi, yoksa tesadüfi olarak mı oluştuğunu belirlemek için bir test yapılabilir.

Hipotez testlerinde yapılacak ilk işlem karar alınacak olayın yapısına göre hipotezleri ortaya koymaktır. Hipotez, anakütle ile ilgili varsayım veya ortaya konan bilgi olarak tanımlanabilir. Bir testte biri temel, diğeri karşıt veya alternatif olmak üzere iki hipotez tanımlanır. Temel veya sıfır hipotez H_0, karşıt veya alternatif hipotez ise H_1 hipotezi olarak ifade edilecektir. Hipotezler oluşturulurken temel hipotez anakütle ile ilgili varsayımın doğruluğunu ifade ederken, karşıt hipotez varsayımın doğru olmadığını ifade edecektir. Test sonucu bu hipotezlerden biri kabul edilirken diğeri reddedilecektir.

Temel hipotezin kabulü anakütle ile ilgili varsayımın doğruluğunu, diğer bir ifade ile varsayım ile örnekten tahmin edilen parametre değeri arasındaki farkın istatistiksel olarak anlamsız olduğunu ortaya koyacaktır. Karşıt hipotezin kabülü ise söylenenlerin aksini ifade edecektir.

Hipotezlerden birinin kabul edilmesi gerekirken, diğeri kabul edilirse hata yapılmış olunacaktır. H_0 hipotezi doğru iken, H_1 hipotezi kabul edilerek, H_0 hipotezi reddedilirse yapılan hata “ I. tip hata”; H_0 hipotezi yanlış iken, H_1 hipotezi yerine H_0 hipotezi kabul edilirse yapılan hata “ II. tip hata ” olarak adlandırılır.

	Ho Kabul	**Ho Red**
Ho Doğru	Doğru Karar	I.Tip Hata
Ho Yanlış	II.Tip Hata	Doğru Karar

Birinci tip hataya α hatası, ikinci tip hataya ise β hatası adı da verilmektedir.

11.2. HİPOTEZ TESTİNİN AŞAMALARI

Sözedildiği gibi hipotez testi bir dizi işlemler sonucu karar verilmesi için yapılmaktadır. Test edilecek parametrelere göre farklı işlemler uygulansa da, test işleminin temel yapısı aynıdır ve genelleştirilebilir. Dört aşamada genelleştireceğimiz test işleminin aşamaları aşağıda açıklanmıştır.

-Hipotezlerin Oluşturulması – Tek ve Çift Taraflı Testler

Test edilecek anakütle parametresi θ ve bunun ile ilgili varsayımı θ_0 ile ifade edelim. θ anakütle ortalaması veya oranı ise θ_0 bunun ile ilgili varsayılan değeri ifade edecektir. Bu anakütleden alınacak n birimli örnek ile tahmin edilen değer ise $\hat{\theta}$ olacaktır. $\hat{\theta}$, θ'nın tahmincisidir ve burada sayısal değeri ifade ettiği varsayılmıştır. Sonuçta $\hat{\theta}=\theta_0$ ise teste gerek yoktur. Test yapılsa bile sonuçta temel hipotez kabul edilecektir. Fakat $\hat{\theta}\neq\theta_0$ olması durumunda aradaki farkın istatistiksel olarak anlamlı olup olmadığı test edilmelidir. Bu durumda temel hipotez,

$H_0 : \theta = \theta_0$

şeklinde oluşturulacaktır.

Karşıt hipotezde temel hipotez gibi farklı bir sayısal değere eşitliği ifade edecek şekilde oluşturulabilir. Bu tür hipotezlere basit hipotez adı verilmektedir. Burada açıklayacağımız gibi eşitsizliklerle oluşturulan hipotezlere ise karmaşık hipotez denilmektedir.

Karşıt hipotez iki farklı şekilde oluşturulabilir. İlk olarak karşıt hipotez eşitsizlik şeklinde kurulur ki, buna çift taraflı hipotez adı verilmektedir.

$H_1 : \theta \neq \theta_0$

Bu tür hipotez testlerine de çift taraflı hipotez testi denmektedir. Anlaşılacağı gibi "çift taraf" küçüklük ve büyüklük ayrımı yapılmadan ikisinin birlikte testinden kaynaklanmaktadır.

Karşıt hipotez büyüklüğü veya küçüklüğü test edecek şekilde de oluşturulabilir. Bu durumda alternatif hipotez tek taraflı hipotez olarak adlandırılır.

$H_1 : \theta > \theta_0$

veya

$H_1 : \theta < \theta_0$

olarak kurulur ve bu tür hipotez testlerine tek taraflı hipotez testleri denmektedir. Karşıt hipotezin büyüklüğü ifade etmesi durumuna sağ taraflı test, küçüklüğü ifade etmesi durumuna ise sol taraflı test adı da verilmektedir.

Çift taraflı testlerde $\hat{\theta}$ ile θ_0 arasındaki ilişki önemli değilken, tek taraflı testlerde önemlidir. $\hat{\theta} > \theta_0$ olması durumunda alternatif hipotezin $\hat{\theta} < \theta_0$ şeklinde oluşturulması anlamsızdır. Yani, büyüklük şeklinde hipotez ancak $\hat{\theta} > \theta_0$ ise kurulmalıdır. Küçüklük içinde aynı açıklamalar geçerlidir.

Aralık tahmininde iki ortalama veya iki oran farkının olduğu gibi birbirinden farklı iki anakütle parametrelerinin aynı olup olmadığı da test edilebilir. Birinci anakütlenin test edilecek parametresi θ_1, ikinci anakütlenin ki θ_2 ise, burada bir varsayım olmayacak, fakat ana kütlelerden alınacak n_1 ve n_2 birimli örneklerden elde edilecek $\hat{\theta}_1$ ve $\hat{\theta}_2$ tahmincilerinin sayısal değerleri ile test yapılacaktır. Nokta tahminin $\theta_1 = \hat{\theta}_1$ ve $\theta_2 = \hat{\theta}_2$ olarak yapılacağı unutulmamalıdır. Bu durumda temel hipotez,

H_0: $\theta_1 = \theta_2$ veya $\theta_1 - \theta_2 = 0$

şeklinde oluşturulur. Alternatif hipotez ise çift taraflı test için,

H_1: $\theta_1 \neq \theta_2$ veya $\theta_1 - \theta_2 \neq 0$

ve tek taraflı test için,

H_1: $\theta_1 > \theta_2$ veya $\theta_1 - \theta_2 > 0$

veya

H_1: $\theta_1 < \theta_2$ veya $\theta_1 - \theta_2 < 0$

şeklinde kurulacaktır. Bu gibi durumlarda alternatif hipotez oluşturulurken $\hat{\theta}_1$ ile $\hat{\theta}_2$ sayısal değerlerinin büyüklüklerine dikkat edilmelidir.

-Hata Payının ve Kritik Değerin Belirlenmesi

Hata payı ve güven olasılığını aralık tahmininden biliyoruz. Hipotez testleri de hata payı ve güven olasılığı ile yapılır. Bu olasılık testin sonucuna güveni ifade etmektedir. Diğer bir ifade ile testin sonucu belirlenen hata payı veya güven olasılığı ile geçerlidir.

Hipotez testlerinde de hata payını α, güven olasılığını da $(1-\alpha)$ olarak ifade edeceğiz. Hipotez testlerinde hata payı α birinci tür hata yapma olasılığıdır ve bu nedenle birinci tür hataya α hatası da denilmektedir. Uygulamada hata payı 0,05 veya 0,01 olarak alınmaktadır. Aslında teorik olarak α veya $(1-\alpha)$ olasılık olduğuna göre testi yapan kişi istediği hata payını seçebilir. Fakat 0,05'ten büyük hata payları sonuçlara güveni sağlamak için uygulamada pek kullanılmamaktadır.

Uygun hata payı belirlendikten sonra testin karar aşamasında kullanılacak tablo değeri de belirlenecektir. Tablo değerine kritik değer adı verilmektedir. Bu değerin belirlenmesi için önce örnekleme dağılımı bilinmelidir. Bu yapılacak testin türüne göre normal dağılımı t dağılımı, F dağılımı veya ki-kare dağılımı olabilir. Bu bölümde yapacağımız açıklamalarda ki-kare dağılımı yer almayacaktır. Örnekleme dağılımına göre düzenlenmiş tablodan seçilen hata payı (α) ve varsa dağılımın serbestlik derecesi de dikkate alınarak kritik değer belirlenir.

-Test İstatistiğinin Hesaplanması

Hipotez testlerinde üçüncü olarak test istatistiği hesaplanır. Test istatistiği test edilecek parametreye ve test edilecek parametre sayısına göre değişir. Örneğin, bir anakütle ortalamasının belirli bir değere eşitliği veya iki anakütle ortalamasının eşitliği test edilirken farklı test istatistikleri kullanılacaktır. Özellikle ortalama ve oranların testinde, örnekleme dağılımı ile ilgili açıklamalarda belirtilen standart hatanın hesaplanması ile ilgili özellikler önemlidir. Standart hatanın doğru olarak hesaplanmasına dikkat edilmelidir.

-Karar

Yapılan test işleminin son adımı hipotezlerden birinin kabul, diğerinin reddedildiği karar aşamasıdır. Kararın verilmesinde hesaplanan test istatistiği ile tablodan bulunan kritik değerler karşılaştırılır. Dağılımın türüne göre kabul ve red bölgeleri oluşturulabilir. Gerekli açıklama daha sonra yapılacaktır. Genelleyerek, test istatistiğinin mutlak değer olarak teorik değerden küçük olması durumunda H_0 hipotezinin, büyük olması durumunda ise H_1 hipotezinin kabul edileceğini söyleyebiliriz.

11.3. ORTALAMALARIN TESTİ

Anakütle ortalaması μ ile ilgili olarak varolan bilgi veya varsayımın doğruluğu, örnekten elde edilecek bilgi ile incelenebilir. Anakütle ortalaması ile ilgili varsayım μ_0 gibi sayısal bir değer ise seçilecek n birimli örnek ortalaması anakütle ortalamasının en iyi tahmincisi olduğundan $\mu = \overline{X}$ olarak tahmin edilecektir. Anakütle ile ilgili varsayımın doğruluğu durumunda $\mu_0 = \overline{X}$ olmalıdır. $\mu_0 \neq \overline{X}$ olması durumunda anakütle ile ilgili varsayımın doğru olmadığı söylenemez. $\overline{X}$ örnekten elde edildiğinden örnekleme hatası söz konusu olabilir. Bu nedenle yapılacak bir test ile aradaki farkın istatistiksel olarak anlamlı olup olmadığı, yani örneklemeden kaynaklanıp kaynaklanmadığı belirlenecektir.

Test için ilk olarak hipotezler belirlenecektir. Yapılacak test çift taraflı da, tek taraflı da olsa temel hipotez,

$$H_0 : \mu = \mu_0$$

şeklinde oluşturulacaktır. Karşıt hipotez ise testin çift veya tek taraflı olmasına göre kurulur. Test çift taraflı ise,

$$H_1 : \mu \neq \mu_0$$

test tek taraflı ise, büyüklük ve küçüklüğe göre,

$$H_1 : \mu < \mu_0$$

veya

$$H_1 : \mu > \mu_0$$

olacaktır.

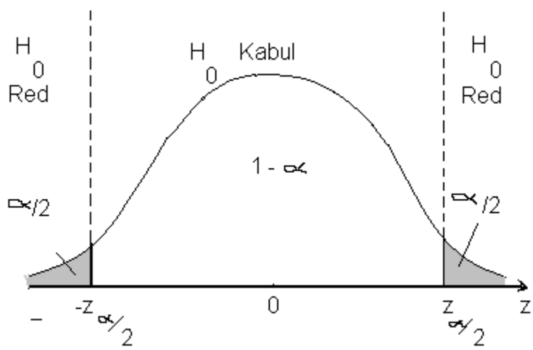
Hipotezler oluşturulduktan sonra α hata payı veya $(1-\alpha)$ güven olasılığı belirlenerek kritik değer bulunacaktır. Kritik değerin belirlenmesinde daha önce ortalamaların aralık tahmininde sözedilen örnekleme dağılımı ile ilgili özellikler dikkate alınır. Anakütle dağılımı normal ise çift taraflı test için $Z_{\alpha/2}$, tek taraflı test için Z_{α} değeri normal dağılım tablosundan bulunacaktır. Anakütle dağılımı normal dağılım kabul ediliyorsa ve örnek birim sayısı 30'dan küçükse örnekleme dağılımı normal dağılım yerine (n-1) serbestlik dereceli t dağılımı olacaktır. Bu nedenle $Z_{\alpha/2}$ ve Z_{α} değerleri yerine t dağılımı tablosundan n-1 serbestlik derecesi ile testin çift veya tek taraflı olmasına göre $t_{\alpha/2,n-1}$ veya $t_{\alpha,n-1}$ değerleri bulunacaktır.

Ortalamaların testinde test istatistiği,

$$Z = \frac{\overline{X} - \mu}{\sigma_{\overline{X}}}$$

olarak hesaplanır. Burada $\overline{X}$ örnekten hesaplanırken μ yerine varsayılan değeri μ_0 kullanılacaktır. Burada dikkat edilmesi gereken standart hatanın hesaplanmasında hata yapılmamasıdır. Standart hatanın seçimin iadeli veya iadesiz olmasına, örnekleme oranına ve anakütle varyansının bilinip bilinmemesine göre nasıl belirlendiği ortalamaların aralık tahmini bölümünde açıklanmıştı. Bu açıklamalara göre standart hata belirlenerek test istatistiği hesaplanmış olacaktır.

Test istatistiği de belirlendikten sonra hipotezlerden hangisinin kabul edileceğine karar verilecektir. Çift taraflı test için H_0 hipotezinin red ve kabul bölgeleri aşağıdaki şekilde gösterilmiştir. H_0 hipotezinin kabulü H_1 hipotezinin reddini, H_0 hipotezinin reddi ise H_1 hipotezinin kabulünü ifade etmektedir.

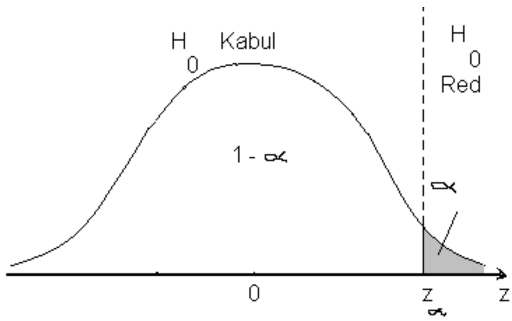


Şekil 10.1

Tek taraflı testte karşıt hipotez,

$H_1 : \mu > \mu_0$

şeklinde oluşturulmuş ise red ve kabul bölgeleri şöyle olacaktır.

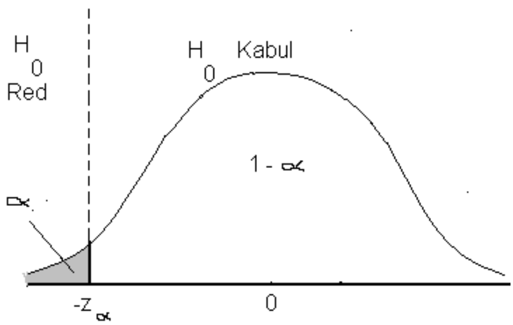


Şekil 10.2

Tek taraflı testte karşıt hipotez,

$H_1 : \mu < \mu_0$

şeklinde oluşturulmuş ise red ve kabul bölgeleri şöyle olacaktır.



Şekil 10.3

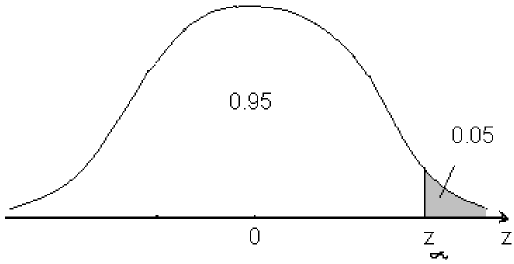
Test istatistiğinin şekillerde düştüğü bölgeye göre hipotezlerden biri kabul diğeri reddedilecektir.

ÖRNEK: Bir makinenin ürettiği A mallarının ağırlıklarının dağılımlarının ortalaması 5 kg. ve standart sapması 1,8 kg.'dır. Makine ayarını kontrol etmek amacı ile 81 birimlik bir örnek alınmış ve örnek ortalaması 5,4 kg. bulunmuştur. Üretilen mamüllerin ortalamasının 5 kg.'dan fazla olduğu söylenebilir mi, 0,05 anlam seviyesi ile inceleyiniz.

$\mu = 5$
$\sigma = 1,8$
$n = 81$
$\overline{X} = 5,4$
$\alpha = 0,05$

$H_0 : \mu = 5$
$H_1 : \mu > 5$

$Z_\alpha \rightarrow \alpha = 0,05$ için



1-0,05 =0,95

0,95 – 0,5 = 0,4500 → Tablodan $Z_\alpha = 1,64$

$$Z = \left| \frac{\overline{X} - \mu}{\sigma_{\overline{X}}} \right| = \left| \frac{\overline{X} - \mu}{\frac{\sigma}{\sqrt{n}}} \right| = \left| \frac{5,4 - 5}{\frac{1,8}{\sqrt{81}}} \right| = \left| \frac{0,4}{0,2} \right| = 2$$

$|Z| > |Z_\alpha|$; 2 >1,64 olduğundan H_0 reddedilerek, H_1 kabul edilir. 0,05 anlam seviyesi ile makinenin ürettiği mamüllerin ortalama ağırlığı 5 kg.'dan fazladır.

11.4. ORANLARIN TESTİ

Anakütle ortalamasında olduğu gibi, anakütle oranı ile ilgili varsayım veya bilgide olabilir. Anakütle oranı P ile ilgili varsayım P_0 ise n birimli örnekten hesaplanacak örnek oranı ise p ise $p=P_0$ olmaması durumunda aradaki farkın istatistiksel olarak anlamlı olup olmadığı test edilecektir. Bu durumda test tek taraflı da çift taraflı da olsa temel hipotez,

$H_0 : P = P_0$

olarak karşıt hipotez çift taraflı test için,

$$H_1 : P \neq P_0$$

tek taraflı test için,

$$H_1 : P < P_0$$

veya

$$H_1 : P > P_0$$

olarak oluşturulacaktır.

Kritik değer ortalamaların testinde açıklandığı gibi belirlenecektir. Küçük örnekler için oranların testi ile uygulamada pek sık karşılaşılmamaktadır. Bu nedenle oranların testinde t dağılımı hemen hemen hiç kullanılmaz.

Oranların testi için test istatistiği,

$$Z = \frac{p - P}{\sigma_p}$$

dir. Burada P yerine P_0 değeri kullanılır. Standart hata σ_p daha önce açıklandığı gibi hesaplanacaktır. Standart hatanın

$$\sigma_p = \sqrt{\frac{pq}{n}}$$

olduğu düşünülürse burada oran olarak biri varsayım, diğeri örnek oranı olmak üzere iki oran bulunmaktadır. Standart hata hesaplanırken oran olarak anakütle ile ilgili varsayım alınacaktır. Hipotezlerden hangisinin kabul edileceği kararı daha önce açıklandığı gibi verilecektir.

ÖRNEK: Bir bölgede yaşayan ailelerin 0,70'inin A malı tüketicisi olduğu kabul edilmektedir. Bu bölgeden 150 aileden oluşan bir örnek alınmış ve bu ailelerin 0,65'inin A malı tüketicisi olduğu anlaşılmıştır. 0,01 anlam seviyesine göre bu bölgede yaşayan ailelerden 0,70'ten daha azının A malı tüketicisi olduğu söylenir mi?

$$P = 0{,}70$$

$$p = 0{,}65$$

$$n = 150$$

$$\alpha = 0{,}01$$

$$H_0 : P = 0{,}70 \qquad (P = P_0)$$

$$H_1 : P < 0{,}70 \qquad (P < P_0)$$

$Z_\alpha = 2{,}33$

$$Z = \left| \frac{p - P}{\sigma_p} \right| = \left| \frac{p - P}{\sqrt{\frac{pq}{n}}} \right| = \left| \frac{0{,}65 - 0{,}70}{\sqrt{\frac{0{,}70 \cdot 0{,}30}{150}}} \right| = \frac{0{,}05}{0{,}038} = 1{,}31$$

$|Z| < |Z_\alpha|$; 1,31<2,33 olduğundan H_0 hipotezi kabul edilir. Bu bölgede yaşayan ailelerden 0,70'ten azının A malı tüketicisi olduğunu söylemek doğru olmayacaktır.

11.5. ORTALAMALAR ARASINDAKİ FARKLARIN TESTİ

Ortalamalar arasındaki farkların aralık tahmini açıklanırken örneklerin bağımsızlığından sözedilmişti. Ortalamalar arasındaki farkların testinde de yapılacak işlemler örneklerin bağımsız veya bağımlı olmasına göre değişmektedir. Bu nedenle konu iki alt başlık altında açıklanacaktır.

11.5.1. Bağımsız Örnekler

İki farklı anakütleden alınan n_1 ve n_2 birimli iki bağımsız örneğin ortalamaları $\overline{X}_1$ ve $\overline{X}_2$ olsun. Örneklerin alındığı anakütlelerin ortalamaları μ_1 ve μ_2 örnek ortalamalarına eşit olarak tahmin edilecektir. Ortalamalar arasında fark olması durumunda, farkın örnekleme hatasından kaynaklanıp kaynaklanmadığının araştırılması istenebilir. Diğer bir ifade ile istatistiksel olarak anakütle ortalamaları birbirine eşit kabul edilebilir mi, ortalamalar arasındaki fark istatistiksel olarak anlamsız mıdır sorusuna yapılacak bir test ile cevap verilebilir.

Ortalamalar arasındaki farkların testinde temel hipotez,

H_0: $\mu_1 = \mu_2$ veya H_0: $\mu_1 - \mu_2 = 0$

şeklinde oluşturulur. Karşıt hipotez ise çift taraflı test için,

H_1: $\mu_1 \neq \mu_2$ veya H_1: $\mu_1 - \mu_2 \neq 0$

tek taraflı testlerde ise,

H_1: $\mu_1 > \mu_2$ veya H_1: $\mu_1 - \mu_2 > 0$

veya

H_1: $\mu_1 < \mu_2$ veya H_1: $\mu_1 - \mu_2 < 0$

olarak oluşturulur.

Kritik değerler daha önce açıklandığı gibi belirlenen hata payına ve testin tek veya çift taraflı olmasına göre normal dağılım tablolarından veya (n_1+n_2-2) serbestlik derecesi ile t dağılımı tablolarından bulunacaktır.

Bağımsız örnekler için ortalama farklarının testinde test istatistiği,

$$Z = \frac{(\overline{X}_1 - \overline{X}_2) - (\mu_1 - \mu_2)}{\sigma_{\overline{X}_1 - \overline{X}_2}}$$

dır. Temel hipotez ortalamalar arasında fark olmadığını ifade ettiğinden $\mu_1 - \mu_2 = 0$ alınırsa test istatistiği,

$$Z = \frac{\overline{X}_1 - \overline{X}_2}{\sigma_{\overline{X}_1 - \overline{X}_2}}$$

olacaktır. Ortalamalar arasındaki farkların standart hatası daha önce açıklandığı gibi varyanslarla ilgili bilgiler ve örnek birim sayıları da dikkate alınarak belirlenecektir. Hipotezlerden hangisinin kabul edileceği kararı da daha önce açıklandığı gibi tablolardan bulunan değerler ile test istatistiği karşılaştırılarak verilecektir.

ÖRNEK: A ve B bölgelerinde yaşayan ailelerin aylık kişi başına beslenme harcamalarını belirlemek amacı ile, A bölgelerinde 140 ve B bölgesinden 90 aileden oluşan iki örnek alınmıştır. A bölgesinden alınan örnekten $\overline{X}_1$=76 milyon TL. ve S_1=7 milyon TL.; B bölgesinden alınan örnekten $\overline{X}_2$ =81 milyon TL. ve S_2=8 milyon TL. bulunmuştur. 0,05 anlam seviyesine göre A ve B bölgelerinde yaşayan ailelerin günlük kişi başına beslenme harcamaları arasında fark olup olmadığını inceleyiniz.

$$\overline{X}_1 = 76 \quad ; \quad S_1 = 7 \quad ; \quad n_1 = 140$$
$$\overline{X}_2 = 81 \quad ; \quad S_2 = 8 \quad ; \quad n_2 = 90$$

1- $H_0 : \mu_1 = \mu_2$
$H_1 : \mu_1 \neq \mu_2$

2- $Z_{\alpha/2}$ =1,96

3-

$$Z = \frac{\overline{X}_1 - \overline{X}_2}{\sigma_{\overline{X}_1 - \overline{X}_2}} = \frac{\overline{X}_1 - \overline{X}_2}{\sqrt{\frac{S_1^2}{n_1} + \frac{S_2^2}{n_2}}} = \frac{76 - 81}{\sqrt{\frac{7^2}{140} + \frac{8^2}{90}}}$$

$$= \frac{-5}{1,030} = -4,85$$

4- $|Z| > |Z_{\alpha/2}|$; 4,85 > 1,96 olduğundan H_1 hipotezi kabul edilir. Harcamalar arasında fark vardır.

ÖRNEK: Önceki örnekte n_1=14 ve n_2=9 kabul ederek örneği tekrar çözünüz.

1- $H_0 : \mu_1 = \mu_2$
$H_1 : \mu_1 \neq \mu_2$

2- n_1 ve n_2 küçük örnekler olduğundan n_1+n_2-2 =14+9-2 =21 serbestlik dereceli t dağılımı $t_{0,05/2,21}$ =2,080

3- $Z = \dfrac{\overline{X}_1 - \overline{X}_2}{\sigma_{\overline{X}_1 - \overline{X}_2}}$

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

$$= \frac{(14-1)(49) + (9-1)(64)}{14 + 9 - 2} = 54{,}7143$$

$$S = 7{,}3969$$

$$\sigma_{\overline{X}_1 - \overline{X}_2} = S\sqrt{\frac{1}{n_1} + \frac{1}{n_2}}$$

$$= 7{,}3969\sqrt{\frac{1}{14} + \frac{1}{9}} = 3{,}1603$$

$$Z = \frac{76 - 81}{3{,}1603} = -1{,}58$$

4- $|1{,}58| < 2{,}080$ olduğundan temel hipotez kabul edilir.

11.5.2. Bağımlı Örnekler

Bazı durumlarda aynı birimlerde yapılan farklı ölçümlerin ortalamalarının arasındaki farkın istatistiksel olarak anlamlılığı da test edilmek istenebilir. Örneğin bir grup öğrenciye uygulanan eğitim öncesi ve sonrası sınavların ortalamaları, hastalara uygulanan tedavi öncesi ve sonrası ölçümlerin ortalaması, makinenin üretiminin ayarlama öncesi ve sonrası ölçümlerin ortalaması gibi. Bu gibi durumlarda bağımlı örnekler sözkonusu olmaktadır.

Aynı birimde yapılan iki farklı ölçümün ortalamaları arasındaki fark, ölçümlerin birbirinden farklarının ortalamasına eşit olacaktır. Ölçümleri X_1 ve X_2 değişkenleri ile gösterelim. Ölçüm farklarını D_i ile ifade edersek,

$$D_i = X_{1i} - X_{2i}$$

olacaktır. Değişkenlerin ayrı ayrı ortalamaları $\overline{X}_1$ ve $\overline{X}_2$ arasındaki fark $\overline{D}$ 'ya eşittir.

$$\overline{X}_1 - \overline{X}_2 = \overline{D}$$

Bu durumda standart değişken

$$Z = \frac{\overline{D} - D}{\sigma_{\overline{D}}}$$

ve standart hata ise,

$$\sigma_D = \sqrt{\frac{n\Sigma D_i^2 - (\Sigma D_i)^2}{n^2(n-1)}}$$

olacaktır. Burada n ölçüm yapılan birim sayısını göstermektedir. İki farklı ölçüm için birim sayısının iki katı alınmamaktadır.

Yukarıda söz edildiği gibi bağımlı örnekler söz konusu olduğunda da hipotezler bağımsız örneklerin ortalamalarının farklarında olduğu gibi çift veya tek taraflı olarak oluşturulur. Kısaca hipotezler değişmemektedir.

Kritik değerin belirlenmesi örnekleme dağılımı ile ilgilidir. Burada da örnekleme dağılımı normal dağılım ve t dağılımı olmaktadır. n <30 olması durumunda (n-1) serbestlik dereceli t dağılımı sözkonusudur. α hata payı ile ilgili tablolardan kritik değerler daha önce açıklandığı gibi belirlenir.

$$Z = \frac{\overline{D}}{S_{\overline{D}}}$$

olacaktır. D'lerin ortalaması,

$$\overline{D} = \frac{\Sigma D_i}{n}$$

ve $\sigma_{\overline{D}}$ değerleri ile test istatistiği,

$$Z = \frac{\frac{\Sigma D}{n}}{\sqrt{\frac{n\Sigma D_i^2 - (\Sigma D_i)^2}{n^2(n-1)}}} = \frac{\Sigma D_i}{\sqrt{\frac{n\Sigma D_i^2 - (\Sigma D_i)^2}{n-1}}}$$

olacaktır. Karar ise daha önce olduğu gibi verilecektir.

ÖRNEK: 9 öğrenci tesadüfi olarak seçilerek 2 hafta süre ile eğitim verilmiştir. Eğitim öncesi ve sonrası yapılan sınav notları aşağıda verilmiştir.

Öğrenci No	1. Sınav	2. Sınav
1	42	75
2	61	60
3	96	90
4	38	57
5	71	60
6	17	35
7	47	63

8	50	94
9	41	88

0,05 hata payı ile öğrencilere verilen eğitim etkili olmuş mudur, inceleyiniz.

Öğrenci No	1. Sınav	2. Sınav	$D_i = X_{1i} - X_{2i}$	D_i^2
1	42	75	-33	1089
2	61	60	1	1
3	96	90	6	36
4	38	57	-19	361
5	71	60	11	121
6	17	35	-18	324
7	47	63	-16	256
8	50	94	-44	1936
9	41	88	-47	2209
			-159	6333

1- $H_0 : \mu_1 = \mu_2$
$H_1 : \mu_1 < \mu_2$

2- n <30

$\alpha = 0,05$

SD = n-1 = 9-1= 8 t dağılımı

$t_{0,05;\,8} = 1,860$

3-

$$Z = \frac{\Sigma D_i}{\sqrt{\dfrac{n\Sigma D_i^2 - (\Sigma D_i)^2}{n-1}}}$$

$$= \frac{-159}{\sqrt{\dfrac{9(6333) - (-159)^2}{9-1}}} = -2,525$$

4- $|Z| > |t_{0,05,8}|$; $|2,525| > |1,860|$ olduğundan H_0 hipotezi reddedilerek, H_1 hipotezi kabul edilir. Buna göre eğitim etkili olmuştur.

11.6. ORANLAR ARASINDAKİ FARKLARIN TESTİ

İki farklı anakütleden alınan bağımsız örneklerin ortalamaları arasındaki farkın anlamlılığı gibi, oranlar arasındaki farkın anlamlılığı da test edilebilir. Anakütle oranı P_1 ve P_2, alınacak n_1 ve n_2 birimli örneklerin oranları p_1 ve p_2 ise anakütle oranları örnek oranlarına eşit olarak tahmin edilecektir. Yapılacak test ile p_1 ve p_2 arasındaki farkın istatistiksel anlamlılığı incelenecektir.

Oranlar arasındaki farkların testinde temel hipotez,

$$H_0 : P_1 = P_2 \quad veya \quad P_1 - P_2 = 0$$

karşıt hipotez ise çift taraflı test için,

$$H_1 : P_1 \neq P_2 \quad veya \quad P_1 - P_2 \neq 0$$

tek taraflı test için ise,

$$H_1 : P_1 > P_2 \quad veya \quad P_1 - P_2 > 0$$

veya

$$H_1 : P_1 < P_2 \quad veya \quad P_1 - P_2 < 0$$

şeklinde oluşturulur.

Oran farklarının testinde kritik değer daha önce açıklandığı gibi belirlenen hata payı veya güven olasılığı ile bulunur. Burada test istatistiği,

$$Z = \frac{(p_1 - p_2) - (P_1 - P_2)}{\sigma_{p_1 - p_2}}$$

dir. Standart hata ise oran farklarının aralık tahmininde açıklandığı gibi belirlenir. Hipotezlerden birinin kabulü kararı da aynı şekilde verilir.

ÖRNEK: İki ambarda bulunan malların bazılarının hatalı olduğu bilinmektedir. Hata oranlarının karşılaştırılması için iki örnek alınmıştır. Birinci örnek 148 birimden oluşmuştur ve bunlardan 37 adedi hatalıdır. İkinci örnek 125 birimden oluşmuştur ve 25 adedi hatalıdır. 0,05 hata payı ile örneklerin alındığı anakütle oranları arasında fark var mıdır?

1- $H_0 : P_1 = P_2$
$H_1 : P_1 \neq P_2$

2- $Z_{\alpha/2} = 1{,}96$

3- $$Z = \frac{|P_1 - P_2|}{\sigma_{p_1 - p_2}} = \frac{|P_1 - P_2|}{\sqrt{\dfrac{p_1 q_1}{n_1} + \dfrac{p_2 q_2}{n_2}}}$$

$$P_1 = \frac{k_1}{n_1} = \frac{37}{148} = 0{,}25$$

$$P_2 = \frac{k_2}{n_2} = \frac{25}{125} = 0{,}20$$

$$Z = \frac{|0{,}25 - 0{,}20|}{\sqrt{\dfrac{(0{,}25).(0{,}75)}{148} + \dfrac{(0{,}20).(0{,}80)}{125}}} = \frac{|0{,}05|}{0{,}05} = 1$$

4- 1<1,96 olduğundan H_0 hipotezi kabul edilir. Örneklerin alındığı anakütle oranları arasında 0,05 anlam seviyesi ile fark yoktur.

11.7. ANAKÜTLE VARYANSININ TESTİ

Önceki bölümlerde ortalamalar ve oranlar ile ilgili testler açıklanmıştı. Bu parametreler gibi varyansların testi de gerekebilir. Örneğin, belirli özellikte malları üreten üreticiler ürettikleri malların özelliklerinin değişkenliğini test etmek isteyebilirler. Bu gibi durumlarda ortalamaların ve oranların testinde olduğu gibi anakütle varyansı ile ilgili bir bilgi veya varsayım gerekecektir.

Normal dağılmış bir anakütle varyansı (σ^2) ile ilgili varsayılan değer σ_0^2 ise, bu varsayımın doğru olup olmadığını belirlemek için n birimli bir örnek alınarak, hesaplanacak S^2 ile test yapılacaktır. Anakütle varyansının testinde de çift veya tek taraflı test yapılabilir. Her iki durum için de temel hipotez anakütle varyansının varsayılan değere eşit oluğunu ifade edecek şekilde,

$$H_0 : \sigma^2 = \sigma_0^2$$

olarak oluşturulacaktır. Karşıt hipotez ise çift taraflı test için,

$$H_1 : \sigma^2 \neq \sigma_0^2$$

tek taraflı test için ise,

$$H_1 : \sigma^2 > \sigma_0^2$$

veya

$$H_1 : \sigma^2 < \sigma_0^2$$

şeklinde oluşturulacaktır.

Anakütle varyansının testinde dağılım χ^2 dağılımıdır. Bu nedenle belirlenen α hata payı ile çift taraflı test için, $\chi^2_{1-\alpha/2,n-1}$ ve $\chi^2_{\alpha/2,n-1}$ değerleri; tek taraflı test için büyüklük sözkonusu ise $\chi^2_{\alpha,n-1}$, küçüklük sözkonusu ise $\chi^2_{1-\alpha,n-1}$ değerleri ki-kare tablosundan bulunacaktır.

Anakütle varyansının testinde kullanılacak test istatistiğini χ^2 ile ifade edersek,

$$\chi^2 = \frac{(n-1)S^2}{\sigma_0^2}$$

dir. Hesaplanan χ^2 test istatistiği belirlenen tablo değerleri ile karşılaştırılarak hipotezlerden biri kabul edilecektir. Çift taraflı testlerde, $\chi^2_{1-\alpha,n-1} \leq \chi^2 \leq \chi^2_{\alpha/2,n-1}$ ise H_0 hipotezi, aksi sözkonusu ise H_1 hipotezi kabul edilecektir. Tek taraflı testlerde $H_1 : \sigma^2 > \sigma_0^2$ hipotezi için

$\chi^2 \geq \chi^2_{\alpha,n-1}$ ise, $H_1 : \sigma^2 < \sigma_0^2$ hipotezi için $\chi^2 \leq \chi^2_{1-\alpha,n-1}$ ise H_0 hipotezi reddedilerek H_1 hipotezi kabul edilecektir.

ÖRNEK: Normal dağılan bir anakütle varyansının 225 olduğu ileri sürülmektedir. Bu anakütleden 25 birimli bir örnek alınarak S^2=312 bulunmuştur. 0,05 hata payı ile anakütle varyansının 225 birimden büyük olduğu söylenebilir mi, inceleyiniz.

1- $H_0 : \sigma^2 = 225$

$H_1 : \sigma^2 > 225$

2- α =0,05

SD= n-1=25-1=24

$\chi^2_{0,05;24} = 36,415$

3-

$$\chi^2 = \frac{(n-1)S^2}{\sigma_0^2}$$

$$= \frac{(25-1).312}{225} = 33,28$$

4- $\chi^2 < \chi^2_{\alpha,n-1}$; 33,28<36,415 olduğundan H_0 hipotezi kabul edilir. Anakütle varyansı 225 birim olarak kabul edilir.

11.8. VARYANS ORANLARININ TESTİ

Ortalamalar ve oranlar arasındaki farkların testi gibi varyansların farklarının testi de gerekebilir. Örneğin, daha önce açıklandığı gibi ortalamalar arasındaki farklar test edilirken varyansların eşit olduğu ileri sürülebilir. Bu gibi durumlarda ortalamalar arasındaki farktan önce varyansların eşitliği test edilebilir. Burada varyansların farklarının testi yerine varyans oranlarının testinden sözedilmesi test istatistiğinin yapısından kaynaklanmaktadır.

Varyans oranlarının veya varyans farklarının testinde de çift ve tek taraflı test yapılabilir. Her iki test içinde temel hipotez,

$$H_0 : \sigma_1^2 = \sigma_2^2 \quad veya \quad \sigma_1^2 - \sigma_2^2 = 0$$

şeklinde oluşturulur. Karşıt hipotez ise çift taraflı test için,

$$H_1 : \sigma_1^2 \neq \sigma_2^2 \quad veya \quad \sigma_1^2 - \sigma_2^2 \neq 0$$

tek taraflı test için ise,

$$H_1 : \sigma_1^2 > \sigma_2^2 \quad veya \quad \sigma_1^2 - \sigma_2^2 > 0$$

veya

$$H_1 : \sigma_1^2 < \sigma_2^2 \quad veya \quad \sigma_1^2 - \sigma_2^2 < 0$$

şeklinde oluşturulur.

Varyans oranlarının testinde dağılım F dağılımıdır. Bu nedenle belirlenen α hata payı ile (n_1-1) ve (n_2-1) serbestlik dereceleri ile çift taraflı test için $F_{\alpha/2,n_1-1,n_2-1}$ ve $F_{\alpha/2,n_2-1,n_1-1}$; tek taraflı testler için büyüklük sözkonusu ise F_{α,n_1-1,n_2-1}; küçüklük sözkonusu ise F_{α,n_2-1,n_1-1} değerleri F tablosundan bulunur.

Daha önce varyans oranlarının tahmininde açıklandığı gibi,

$$F = \frac{S_1^2 / \sigma_1^2}{S_2^2 / \sigma_2^2}$$

dir. Testte temel hipotez $\sigma_1^2 = \sigma_2^2$ olduğundan test istatistiği

$$F = \frac{S_1^2}{S_2^2}$$

olacaktır.

Test istatistiği (F), S_1 ve S_2'nin büyüklüğüne göre belirlenecektir. Büyük olan S değeri payda yeralacaktır. Çift taraflı test için,

$$S_1^2 \geq S_2^2 \quad ise \quad F = \frac{S_1^2}{S_2^2} \geq F_{\alpha/2,n_1-1,n_2-1}$$

$$S_2^2 \geq S_1^2 \quad ise \quad F = \frac{S_2^2}{S_1^2} \geq F_{\alpha/2,n_2-1,n_1-1}$$

olması durumunda H_1 hipotezi kabul edilecektir. Tek taraflı testlerde

$H_1 : \sigma_1^2 > \sigma_2^2$ için,

$$F = \frac{S_1^2}{S_2^2} \geq F_{\alpha,n_1-1,n_2-1}$$

$H_1 : \sigma_1^2 < \sigma_2^2$ için,

$$F = \frac{S_2^2}{S_1^2} \geq F_{\alpha, n_2-1, n_1-1}$$

olması durumunda H_1 hipotezi kabul edilecektir.

ÖRNEK: Normal dağılmış anakütleden n_1=6 ve n_2=6 birimden oluşan iki örnek alınarak $S_1^2 = 186$ ve $S_2^2 = 141$ olarak belirlenmiştir. 0,05 hata payı ile S_1^2 'nin ve S_2^2 'den büyük olduğu söylenebilir mi, inceleyiniz.

1- $H_0 : \sigma_1^2 = \sigma_2^2$
$H_1 : \sigma_1^2 > \sigma_2^2$

2- α =0,05

SD =n_1-1 =6-1= 5

SD =n_2-1=6-1 =5

$$F_{\alpha, n_2-1, n_1-1} = F_{0,05;5,5} = 5{,}05$$

3- $F = \frac{S_1^2}{S_2^2} = \frac{186}{141} = 1{,}319$

4- 1,319<5,05 olduğundan H_0 hipotezi kabul edilir. Varyanslar birbirine eşittir.

11.9. ÇÖZÜMLÜ ÖRNEKLER

ÖRNEK 1: A marka ampullerin ortalama kullanılma sürelerinin 1100 saat olduğu ileri sürülmektedir. Ampullerin ortalama kullanılma sürelerini belirlemek için 64 ampulden oluşan bir örnek alınarak, örnek ortalaması $\overline{X} = 1093$ saat ve $S^2 = 900$ saat bulunmuştur. 0,01 hata payı ile ampüllerin ortalama kullanılma sürelerinin 1100 saatten az olduğu söylenebilir mi?

Çözüm:

$\mu = 1100$

$S^2 = 900 \rightarrow S = 30$

$n = 64$

$\overline{X} = 1093$

$\alpha = 0{,}01$

1- $H_0 : \mu = 1100$
$H_1 : \mu < 1100$

2- 1-α =0,99

0,99 – 0,5 = 0,49 → Tablodan Z_α = 2,33

3- $Z = \left|\dfrac{\overline{X}-\mu}{\sigma_{\overline{X}}}\right| = \left|\dfrac{\overline{X}-\mu}{\dfrac{S}{\sqrt{n}}}\right| = \left|\dfrac{1093-1100}{\dfrac{30}{\sqrt{64}}}\right| = \left|\dfrac{-7}{3,75}\right| = 1,86$

4- $|Z| < |Z_\alpha|$; 1,86<2,33 olduğundan H_0 kabul, H_1 reddedilir. Ampullerin ortalama kullanılma sürelerinin 0,01 anlam seviyesine göre 1100 saatten az olduğu söylenemez.

ÖRNEK 2: Bir fabrikada üretilen A parçalarının çapları 33 mm.'dir. Üretim kontrolü için fabrika üretiminden 200 birimlik bir örnek alınarak $\overline{X} = 34$ mm. ve S =4 mm. bulunmuştur. Buna göre parçaların çaplarının 33 mm. olduğu kabul edilebilir mi? 0,01 hata payı ile inceleyiniz.

Çözüm:

$\mu = 33$

$S = 4$

$n = 200$

$\overline{X} = 34$

$\alpha = 0,01$

1-

$H_0 : \mu = \overline{X}$

$H_1 : \mu \neq \overline{X}$

2- 1-α =0,99

0,99 – 0,5 = 0,495 → Tablodan Z_α = 2,57

3- $Z = \left|\dfrac{\overline{X}-\mu}{\sigma_{\overline{X}}}\right| = \left|\dfrac{\overline{X}-\mu}{\dfrac{S}{\sqrt{n}}}\right| = \left|\dfrac{34-33}{\dfrac{4}{\sqrt{200}}}\right| = \left|\dfrac{1}{0,28}\right| = 3,57$

4- $|Z| > |Z_{\alpha/2}|$; 3,57>2,57 olduğundan H_0 reddedilir, H_1 kabul edilir. Parçaların çaplarının 33 mm. olduğu kabul edilmez.

ÖRNEK 3: Normal olduğu bilinen ve μ=220 olan bir anakütleden n=25 birimlik bir örnek alınarak $\overline{X} = 216$ ve S=12 bulunmuştur. 0,95 güven olasılığı ile anakütle ortalaması ile örnek ortalaması arasında fark olup olmadığını inceleyiniz.

Çözüm:

1-

$H_0 : \mu = 220$
$H_1 : \mu \neq 220$

2- n <30, anakütle dağılımı normal ve σ bilinmediğinden t- dağılımı tablosundan,

$\alpha/2 = 0{,}025$ ve 25-1 =24 serbestlik derecesi ile $t_{\alpha/2} = 2{,}064$

3- Z yerine t kullanalım.

$$t = \left|\frac{\overline{X}-\mu}{\sigma_{\overline{X}}}\right| = \left|\frac{\overline{X}-\mu}{\frac{S}{\sqrt{n}}}\right| = \left|\frac{216-220}{\frac{12}{\sqrt{25}}}\right| = 1{,}66$$

4- $|t| < |t_{\alpha/2}|$; 1,66<2,064 olduğundan H_0 hipotezi kabul edilir. Örnek ortalaması ile anakütle ortalaması arasında fark yoktur.

ÖRNEK 4:

$\overline{X}_1 = 18\,; S_1 = 3\,; n = 49 \quad ve \quad \overline{X}_2 = 21\,; S_2 = 4\,; n = 36$

olan iki örneğin μ= 20 olan bir anakütleden alındığı ileri sürülmektedir. Bu örneklerin 0,10 hata payı ile bu anakütleden alınıp alınmadığını belirleyiniz.

Çözüm:

I. Örnek İçin;

1-

$H_0 : \mu = 20$
$H_1 : \mu \neq 20$

2- 1-α =0,90

0,90/2 = 0,45 → Tablodan $Z_{\alpha/2} = 1{,}64$

3- $$Z = \left|\frac{\overline{X}_1-\mu}{\sigma_{\overline{X}_1}}\right| = \left|\frac{\overline{X}_1-\mu}{\frac{S_1}{\sqrt{n_1}}}\right| = \left|\frac{18-20}{\frac{3}{\sqrt{49}}}\right| = 4{,}66$$

4- $|Z| > |Z_{\alpha/2}|$; 4,66>1,64 olduğundan H_1 kabul edilir. Anakütle ve örnek ortalaması arasında 0,10 anlam seviyesi ile fark vardır. Bu örnek bu anakütleyi temsil etmemektedir ve bu anakütleden alınmadığı söylenebilir. $(\mu \neq \overline{X})$

II. Örnek İçin;

1- Hipotezler :

$H_0 : \mu = 20$
$H_1 : \mu \neq 20$

2- $Z_{\alpha/2} = 1{,}64$

3- $Z = \left|\dfrac{\overline{X}_2 - \mu}{\sigma_{\overline{X}_2}}\right| = \left|\dfrac{\overline{X}_2 - \mu}{\dfrac{S_2}{\sqrt{n_2}}}\right| = \left|\dfrac{21-20}{\dfrac{4}{\sqrt{36}}}\right| = 1{,}5$

4- $|Z| < |Z_{\alpha/2}|$; 1,5<1,64 olduğundan H_0 kabul edilir. Anakütle ve örnek ortalaması arasında 0,10 anlam seviyesi ile fark yoktur. Bu örnek bu anakütleyi temsil etmektedir ve bu anakütleden alındığı söylenebilir. $(\mu = \overline{X})$

ÖRNEK 5: 2000 birimlik bir anakütlenin ortalamasının μ=140 olduğu ileri sürülmektedir. Bu anakütleden iadesiz seçimle alınan 121 birimlik örnekte $\overline{X} = 144$ ve S=9 bulunmuştur. 0,05 hata payı ile anakütle ortalaması 140 olarak kabul edilebilir mi?

1- $H_0 : \mu = 140$
$H_1 : \mu \neq 140$

2- $Z_{\alpha/2} = 1{,}96$

3 $Z = \dfrac{\overline{X} - \mu}{\sigma_{\overline{X}}}$

$\sigma_{\overline{X}}$ için; $\dfrac{n}{N} = \dfrac{121}{2000} = 0{,}0605 > 0{,}05$ olduğundan

$$\sigma_{\overline{X}} = \frac{S}{\sqrt{n}} . \sqrt{\frac{N-n}{N-1}} = \frac{9}{\sqrt{121}} . \sqrt{\frac{2000-121}{2000-1}} = 0{,}79$$

$$Z = \left|\frac{144-140}{0{,}79}\right| = 5{,}06$$

4- $|Z| > |Z_{\alpha/2}|$; 5,06>1,96 olduğundan H_0 hipotezi reddedilir. Anakütle ortalaması 140 olarak kabul edilemez.

ÖRNEK 6: Bir bölgede yapılan bir araştırma ile marjinal tüketim eğiliminin 0,68 olduğu belirlenmiştir. Aynı bölgeden n= 500 olan bir örnek alınarak marjinal tüketim eğilimi 0,76 olarak belirlenmiştir. 0,99 güven olasılığı ile marjinal tüketim eğilimi 0,68'den fazla mıdır, inceleyiniz.

Çözüm:

$P = 0{,}68$

$p = 0{,}76$

$n = 500$

$\alpha = 0{,}01$

1- $H_0 : P = 0{,}68$
$H_1 : P > 0{,}68$

2- Z_α =2,33

3- $Z = \dfrac{p - P}{\sigma_p}$

$$\sigma_p = \sqrt{\frac{pq}{n}} = \sqrt{\frac{(0{,}68)(0{,}32)}{500}} = 0{,}0209$$

$$Z = \frac{0{,}76 - 0{,}68}{0{,}0209} = 3{,}82$$

4- $|Z| > 3{,}82$ olduğundan H_0 hipotezi reddedilir. Marjinal tüketim eğilimi 0,68'den fazladır.

ÖRNEK 7: Bir fabrikada üretilen mamüllerin 0,15'inin hatalı olduğu ileri sürülmektedir. Hata oranını belirlemek amacı ile 250 mamülden oluşan bir örnek alınmış ve 250 mamülden 46 adedinin hatalı olduğu anlaşılmıştır. 0,05 hata payı ile bu fabrika üretiminin 0,15'inin hatalı olduğunu söylemek doğru mudur?

Çözüm:

1- $H_0 : P = 0{,}15$
$H_1 : P \neq 0{,}15$

2- $Z_{\alpha/2}$ = 1,64

3- $Z = \left|\dfrac{p - P}{\sigma_p}\right| = \left|\dfrac{p - P}{\sqrt{\dfrac{pq}{n}}}\right|$

$$p = \frac{46}{250} = 0{,}184$$

$$Z = \left|\frac{0{,}184 - 0{,}15}{\sqrt{\dfrac{(0{,}15).(0{,}85)}{250}}}\right| = \left|\frac{0{,}034}{0{,}022}\right| = 1{,}54$$

4- $|Z| < |Z_{\alpha/2}|$; 1,54 <1,64 olduğundan H_0 hipotezi kabul edilir. Üretimin 0,05'inin hatalı olduğu söylenebilir.

ÖRNEK 8: İki aynı fabrikada aynı mamül üretilmektedir. Mamüllerin boy uzunluklarını karşılaştırmak için iki fabrika üretiminden birer örnek alınmıştır. Birinci fabrika için $\sigma_1 = 12$ mm. ve ikinci fabrika için $\sigma_2 = 14$ mm. olarak bilinmektedir. Birinci örnek 70 mamülden oluşmuş ve ortalaması 198 mm. ; ikinci örnek 80 mamülden oluşmuş ve ortalaması 201 mm. olarak bulunmuştur. 0,10 hata payı ile fabrikalarda üretilen mamüllerin boy uzunlukları arasında fark olup olmadığını belirleyiniz.

Çözüm:

$\sigma_1 = 12$; $\overline{X}_1 = 198$; $n_1 = 70$

$\sigma_2 = 14$; $\overline{X}_2 = 201$; $n_2 = 80$

1- $H_0 : \mu_1 = \mu_2$
$H_1 : \mu_1 \neq \mu_2$

2- $Z_{\alpha/2} = 1,64$

3- $$Z = \left|\frac{\overline{X}_1 - \overline{X}_2}{\sigma_{\overline{X}_1 - \overline{X}_2}}\right| = \left|\frac{\overline{X}_1 - \overline{X}_2}{\sqrt{\dfrac{\sigma_1^2}{n_1} + \dfrac{\sigma_2^2}{n_2}}}\right|$$

$$= \frac{|198 - 201|}{\sqrt{\dfrac{12^2}{70} + \dfrac{14^2}{80}}} = \frac{|3|}{2,12} = 1,41$$

4- $|Z| < |Z_{\alpha/2}|$; 1,41<1,64 olduğundan H_0 hipotezi kabul edilir. Fabrikalarda üretilen mamüllerin boy uzunlukları arasında fark yoktur.

ÖRNEK 9: A malları 1 kg. lık ambalajlarda piyasaya sürülmektedir. A malının kilogramındaki a maddesi miktarını belirlemek amacı ile bu mallardan biri n_1=10 diğeri n_2=12 birimden oluşan iki örnek alınarak $\overline{X}_1 = 21$ gr. ve S_1 =3 gr.; $\overline{X}_2 = 29$ gr. ve S_2 =4 gr. bulunmuştur. 0,01 hata payı ile ortalamalar arasında fark var mıdır? İnceleyiniz. (a maddesi miktarları dağılımı normal olarak bilinmektedir.)

Çözüm:

1- $H_0 : \mu_1 = \mu_2$
$H_1 : \mu_1 \neq \mu_2$

2- $n_1 + n_2 - 2 = 10+12-2 = 20$; σ_1 ve σ_2 bilinmiyor ve dağılımlar normal olduğundan, $\alpha/2 = 0{,}01/2 = 0{,}005$ ve 20 serbestlik derecesi ile t- dağılımı tablosundan $t_{\alpha/2} = 2{,}845$

3- $$t = \frac{\left|\overline{X}_1 - \overline{X}_2\right|}{\sigma_{\overline{X}_1 - \overline{X}_2}}$$

$$S^2 = \frac{(n_1-1)S_1^2 + (n_2-1)S_2^2}{n_1+n_2-2} = \frac{(10-1).3^2 + (12-1).4^2}{10+12-2} = 12{,}85$$

$$S = \sqrt{12{,}85} = 3{,}58$$

$$\sigma_{\overline{X}_1 - \overline{X}_2} = S\sqrt{\frac{1}{n_1} + \frac{1}{n_2}} = 3{,}58\sqrt{\frac{1}{10} + \frac{1}{12}} = 1{,}53$$

$$t = \frac{21-29}{1{,}53} = 5{,}22$$

4- $|t| > |t_{\alpha/2}|$; 5,22 >2,845 olduğundan H_1 hipotezi kabul edilir. Örnek ortalamaları arasında fark vardır.

ÖRNEK 10: İki iş kolunda çalışan işçilerin saat ücretlerini karşılaştırmak için iş kollarından birer örnek alınmıştır. Birinci iş kolu için $n_1 = 20$, $\overline{X}_1 = 188$ TL., $S_1^2 = 55$ TL.; ikinci iş kolunda $n_2 = 22$, $\overline{X}_2 = 5$ TL., $S_2^2 = 6$ TL. bulunmuştur. 0,05 hata payı ile birinci iş kolunda çalışan işçilerin ikinci iş kolunda çalışanlardan daha fazla ücret aldığı söylenebilir mi? İnceleyiniz.

Çözüm:

1- $$H_0 : \mu_1 = \mu_2$$
$$H_1 : \mu_1 \neq \mu_2$$

2- $Z_\alpha = 1{,}64$

3- $$Z = \frac{\overline{X}_1 - \overline{X}_2}{\sigma_{\overline{X}_1 - \overline{X}_2}} = \frac{\overline{X}_1 - \overline{X}_2}{\sqrt{\frac{S_1^2}{n_1} + \frac{S_2^2}{n_2}}} = \frac{8-5}{\sqrt{\frac{5}{20} + \frac{6}{22}}} = \frac{3}{0{,}72} = 4{,}16$$

4- $|Z| > |Z_\alpha|$; 4,16 > 1,64 olduğundan H_1 hipotezi kabul edilir. Birinci iş kolunda işçilerin ikinci iş kolundan daha az ücret aldıkları söylenebilir.

ÖRNEK 11: A ve B bölgelerinde ilkokul çağında olan çocukların okula gidip gitmedikleri araştırılmaktadır. Bu nedenle A bölgesinden 200, B bölgesinden 250 çocuktan oluşan birer örnek alınmış ve A bölgesinde ilkokul çağındaki çocukların

0,71'inin; B bölgesinde 0,62'sinin okula devam ettiği anlaşılmıştır. 0,10 hata payı ile A bölgesinde ilkokul çağındaki çocukların okula devam etme oranının B bölgesinden fazla olduğu söylenebilir mi?

Çözüm:

1- $H_0 : P_A = P_B$
$H_1 : P_A > P_B$

2- $Z_\alpha = 1{,}28$

3-

$$Z = \frac{|P_A - P_B|}{\sigma_{p_A - p_B}} = \frac{|P_A - P_B|}{\sqrt{\frac{p_A q_A}{n_A} + \frac{p_B q_B}{n_B}}}$$

$$= \frac{|0{,}71 - 0{,}62|}{\sqrt{\frac{(0{,}71).(0{,}29)}{200} + \frac{(0{,}62).(0{,}38)}{250}}} = 2{,}04$$

4- $|Z| > |Z_\alpha|$; 2,04 > 1,28 olduğundan H_1 hipotezi kabul edilir. A bölgesinde okul çağında olan çocukların okula devam etme oranları B bölgesinden büyüktür.