Tam Sayım ve Örnekleme

Hakkında araştırma yapılacak birimler topluluğu olan evrenin özelliklere hakkında bilgiler evren veri kümesinin ya da örneklem veri kümesinin çözümlenmesiyle ulaşılabilir.

Tam sayım: Sonlu evrenin bütün birimleri üzerinden verinin derlenmesidir. Elde edilen veri kümesinin çözümlenmesi ile elde edilen sayısal niceliklerse *parametre* olarak adlandırılır.

Örnekleme: Evrenden belirli yöntemlerle seçilmiş olan ve seçildiği evreni temsil ettiği düşünülen birimler kümesine denir.

Örneklemeyi Gerekli Kılan Nedenler

Evren sonsuz ise birim sayısı sınırsızdır dolayısıyla tam sayım mümkün olamayacağından örneklemeye başvurulabilir.

Evren sonlu ise maliyet, zaman, doğru veri elde etme, incelecek birimlerin fiziksel zarara uğraması ve evreni oluşturacak birimlerin değişkenliği kriterleri göz önünde tutularak tam sayım da örnekleme de tercih edilebilir.

Örneklem için Birim Seçme Yöntemleri;

- Keyfi seçim ve
- Rassal seçim olmak üzere ikiye ayrılır.

Keyfi seçim, tüm birimlere belli bir olasılıkla seçilme şansı verilmeyip araştırmacı tarafından seçilecek birimlerin belirlenmesi durumudur.

Rassal seçim, sonlu evrenlerde rassal örnekleme seçimi iki şekilde yapılabilir:

- Kura seçimi,
- Sistematik seçim.

Örnekleme Sürecinin Aşamaları

Genel olarak örnekleme süreci 5 aşamadan oluşur ve kısaca şöyle açıklanabilir:

- **Evrenin tanımlanması:** Gözlem birimi, örnekleme birimi, yer ve zamanın belirlendiği süreçtir. Ayrıca gözlem ile örnekleme birimleri aynı ise yalnızca birim kavramı kullanılır ve evren tanımı yapılırken açıklık, kesinlik, amaca uygunluk ile güçlük oluşturmama gibi ilkelerde göz önünde tutulur.
- **Çerçevenin belirlenmesi:** Çerçeve sonlu evrenin tüm birimlerini içeren liste, tablo veya cetveldir. Ayrıca sonsuz evrenlerin örneklendirilmesinde çerçeve aşaması yoktur. Güncel çerçevelerin kullanımı tam sayım içinde örnekleme içinde gereklidir ancak güncel olmayan çerçeve kullanımları da maliyet ve zaman göz önünde tutularak kabul edilebilir.
- **Örnekleme yönteminin seçilmesi:** *Keyfi* ya da *rassal* olmak üzere ikiye ayrılan bu yöntemler örnekleme birimlerinin belirlenmesini sağlar.
- **Örnekleme hacminin belirlenmesi:** Örnekleme hacmi "n" sembolü ile gösterilir ve örneklemenin sayısını ifade eder. Bu sayının değeri nitel değerlendirmeler ve nicel yöntemler yardımı ile yapılır. Nitel değerlendirmelerde esas olan faktörler,
  - Evrenin homojenliği,
  - Araştırmada verilecek kararın önemi,
  - Araştırmanın yapısıdır.
- **Örneklemenin Seçimi:** Keyfi ya da rassal olan örnekleme yöntemlerinin kullanılmasıyla belirlenen birimlerden veriler derlenir.

Örnekleme Yöntemleri

Olasılıklı ve **olasılıklı olmayan** yöntemler olmak üzere ikiye ayrılır.

Olasılıklı Olmayan Örnekleme Yöntemleri: Birim seçiminin uygulayıcıların istek ve değer yargılarına bağlı olarak keyfi yapıldığı örnekleme süreçleridir. Olasılıklı olmayan bu yöntemle, tüm birimlere belli olasılıkla seçilme şansının verilmediği söylenebilir ancak bu durum örneklemin evreni temsil edemeyeceği anlamına gelmez.

Bu yöntemlerde örneklem için birim seçimi, örneklemin hacmi keyfi olup bulunan istatistikler evren parametreleri için genellenemez. Dörde ayrılan bu yöntemler kısaca şöyle açıklanabilir:

- **Kolayda örnekleme,** doğru zaman ve yerde araştırma konusu ile ilişkili gönüllü birimlerle oluşturulan örnekleme yöntemdir. Zaman ve maliyet açısından ekonomikliğinden ancak temsil yönünden yanlılığından söz edilebilir.
- **Yargısal örnekleme,** temsili örneklem oluşturacağı düşünülen birimlerin kriterlere dayalı olarak evrenden seçilmesine dayalı olan bir tür kolayda örnekleme metodudur. Ancak uzman görüşlerinin var olmasıyla kolayda örneklemeden ayrışır.
- **Kota örneklemesi,** sonlu evrenin öncül bilgiler yardımı ile heterojen birimlerden oluştuğu biliniyorsa her tabakadan birim sayıları ile orantılı olarak seçilen birimlerce örneklem oluşturulur.
- **Kartopu örneklemesi,** rassal olarak bir birey belirlenir ve tanıdığı bireye de ulaşılır. Referanslara göre bu süreç örnekleme hacmine ulaşılana kadar devam ettirilir.

Olasılıklı olmayan örnekleme yöntemlerinde birimlerin keyfi olması tek yönlü hatalara sebep olur. Bu aşamada olasılıklı örnekleme yöntemlerinden bahsedilebilir.

Olasılıklı Örnekleme Yöntemleri: Birimlerin rassal olarak belirlenmesine dayalı örnekleme yöntemleridir. Her birime sıfırdan farklı bir olasılık değeriyle seçilme şansı tanınır.

Basit rassal örnekleme, sonlu evrenlerde basit rassal örnekleme: "N" hacimli evrende, "n" hacimli örneklemde yer alma şansının her birim için eşit tutulduğu örneklem yöntemidir.

Ancak mümkün tüm "n" hacimli örneklemler oluşturulmaz, oluşturulabileceği varsayılır. Yalnızca mümkün bir örneklem oluşturularak çalışılır.

Sonsuz evrenlerde basit rassal örnekleme: Sonsuz evrende tüm sonuçlar listelenemez. Sonsuz evrenin birimleri kararlı ise $x_1, x_2, \ldots, x_n$ gözlem değerleri $X_1, X_2, \ldots, X_n$ rassal olan değişkenlerinin birer gerçekleşmesidir. Bu bağımsız ve benzer olasılık dağılımlı rassal değişkenlerin oluşturduğu topluluk örneklemdir.

Tabakalı örnekleme, Heterojen yapılı evrende birimlerin tabakalandırılması ve her tabakadan rassal olarak yapılan seçimlerle oluşturulan örneklemedir. Tabakalama kriterlerinin belirlenmesi, tabakaların oluşturulması, tabakalardan birimlerin seçilmesi ve verilerin derlenmesi ile bu örnekleme gerçekleştirilir.

Tabaka içi homojenlik tabaka içi varyansı; buda evren tahminleyicisinin varyansını düşürecektir.

Sistematik örnekleme, evrendeki birimler numaralandırılır. Örneklem hacmi belirlenir. Evren hacmi ile örneklem hacmi oranlanarak büyütme faktörü hesaplanır ve bu faktörle 1 arasından bir a sayısı belirlenir. a sayısından başlanarak k artışla a+(n-1)k sıra nolu birime kadar n hacimli sistematik örneklem oluşturulur.

Tek aşamalı ve çok aşamalı küme örneklemesi, genellikle coğrafi kriterlere göre oluşturulan kümeler tabakalı örneklemede olduğu gibi evren birimlerinin gruplarıdır. Tüm kümeler arasından rassal olarak seçilen kümelerin tüm birimleri ile örneklem oluşturulur.

Bir kümeleme ile gözlem birimleri elde ediliyorsa **tek aşamalı;** iki ve üssü kümeleme aşamaları ile elde ediliyorsa **çok aşamalı** kümeleme yapılmış olur.

Evren hacmi büyük, birimler homojen ve çerçeve oluşturmak güç ise bu örnekleme kullanılabilir.

Örnekleme Dağılımları

Parametre (θ) evrene ilişkin sayısal karakteristikleri ifade eder. Tam sayım yapılamadığı durumlarda örnekleme istatistikleri $(\hat{\theta})$ kullanılır. Her n hacimli örneklemden elde edilecek parametre değerleri evren parametresine eşit, evren parametresinden büyük yada küçük olabilir. O halde bir örneklemden hesaplanan istatistiğe değil bu istatistiğin tüm n hacimli örneklemlerde alacağı değerlerin dağılımından yararlanılmalıdır.

Örneklemlerden hesaplanan istatistiklerin olasılık dağılımları *örnekleme dağılımı*dır. Ancak bunların her biri hesaplanmaz. Bunun yerine rassal bir değişken olan örneklem istatistiğinin olasılık dağılımı örneklem dağılımı olarak ele alınır.

Burada örneklem istatistiklerinden $\bar{X}$ ve p'nin örnekleme dağılımları ve özellikleri incelenecektir.

Ortalamanın Örneklem Dağılımı: Rassal bir değişken olan $\bar{X}$, örneklem istatistiklerinden olup örneklem ortalamasıdır. Bu rassal değişkenin olasılık dağılımı ise ortalamanın örnekleme dağılımı olarak adlandırılır. Her $\bar{X}$ değerinin elde edilmesi varsayımı ile oluşturulacak { $\bar{X}_i$} frekans dağılımı elde edilir ve böylelikle ortalamanın örnekleme dağılımı bulunur.

$\bar{X}$ 'nın Dağılımının Özellikleri: Ortalamanın örnek dağılımın ortalaması $\mu = \mu_{\bar{X}}$ ile standart sapması $\sigma_{\bar{X}}$'tir. Ayrıca $\sigma^2_{\bar{X}}$ standart hatanın karesi olup varyanstır.

$\bar{X}$ 'nın dağılımının ortalaması, $\bar{X}$ 'nın örnekleme dağılımın ortalaması $\bar{X}$ 'nın beklenen değeridir.

$$E(\bar{X}) = \mu$$

Burada n arttıkça $\bar{X}$ 'nın örnekleme dağılımının ortalamasının evren ortalamasına yaklaştığı hatta yeterli büyüklükten sonra $\bar{X}$ 'nın örnekleme dağılımının normal dağılıma yaklaştığı söylenebilir.

$\bar{X}$ 'nın dağılımının standart hatası, $\bar{X}$ 'nın standart sapması olarak da adlandırılabilir ve $\sigma_{\bar{X}}$ ile gösterilir. Ortalamanın örnekleme dağılımının değişkenliğinin göstergesidir.

Burada evren değişkenliği σ ile örneklem genişliğine bağlıdır. Örneklem hacmi arttıkça $\sigma_{\bar{X}}$ küçülür ve μ hakkında daha az hatalı sonuçlar elde edilebilir.

Merkezi limit teoremi, örneklem hacmi büyüdükçe evren dağılımına bakılmaksızın $\bar{X}$ 'ın örneklem dağılımın normal dağılıma yaklaştığı söylenebilir. Dağılım ortalaması μ ; varyansı σ^2/n olup n≥30 yeterli örneklem hacmi olarak kabul edilir.

Örnekleme Oranı p'nin Örnekleme Dağılımı: İkili sonuçlar içeren durumlar evrende araştırma konusu ise evrendeki birimlerden araştırılan ikili özellikle ilgili olanların oranları evren oranı olarak isimlendirilir ve $\prod$ ile gösterilir. İlgilenen özelliğe sahip birim R ile gösterilirse evren oranı;

$\prod = R/N$ şeklinde ifade edilir.

İlgilenilmeyen özelliğe sahip birim ise Q ile gösterilirse;

$n/N \leq 0.05$ olarak elde edilir.

Tam sayıp yapılmadığı zaman hesaplanamayacak olan $\prod$ örneklem istatistiklerinden yararlanılarak elde edilebilir.

Bu bağlamda ilgilenilen özellik r olmak üzere n hacimli örneklemde örneklem oranı (p);

$p = r/n$ şeklinde ifade edilir.

İlgilenilmeyen özelliğe sahip birim ise q ile gösterilirse;

$$q = \frac{n-r}{n} = 1 - p$$ olarak elde edilir.

Evrenden yalnızca bir n hacimli örneklem rassal olarak ele alınır ve p oranı hesaplanarak P rassal değişkeninin tüm n hacimli örneklemlerde aldığı değerlerin dağılımına dayalı olarak oranların örneklem dağılımı elde edilir.

Ortalama ve varyans, evren oranı yalnızca n hacimli bir örneklem için bulunan istatistikle değil rassal değişken olan p istatistiğinin örnekleme dağılımının yardımı ile elde edilir.

p oranının beklenen değeri evren oranına eşittir.

$$E(p) = \prod$$

Ayrıca sonsuz evrenlerde ve örnekleme oranı $n/N \leq 0.05$ ise σ^2 'nin bilinmesi halinde varyans;

$$\sigma_p^2 = \frac{\prod(1-\prod)}{n}$$

σ^2 'nin bilinmemesi halindeyse varyans;

$$s_p^2 = \frac{p(1-p)}{n}$$

eşitlikleri yardımı ile elde edilir.

Evren sonsuz ve örnekleme oranı $n/N \leq 0.05$ ise σ^2 'nin bilinmesi halinde varyans;

$$\sigma_p^2 = \frac{\prod(1-\prod)}{n} . \frac{N-n}{N-1}$$

σ^2 'nin bilinmemesi halindeyse varyans;

$$s_p^2 = \frac{p(1-p)}{n} . \frac{N-n}{n}$$

eşitlikleri yardımı ile elde edilir.

Dağılım Şekli ve Merkezi Limit Teoremi, örnekleme planındaki basit rassal örneklemen hacmi (n) arttıkça örneklem oranı p'nin örnekleme dağılımı normal dağılıma yakınsar. Bu bağlamda p rassal değişkeninin standartlaştırılması ile evren hakkında bilgi üretilebilir.

Örneklem hacminin belirlenmesinde nicel yöntemler,

- Karşılanabilecek maliyeti esas alan yöntem ve
- Kabul edilebilir hata düzeyini esas alan yöntem olmak üzere ikiye ayrılır.

İstatistiksel Tahminleme

Rassal bir örneklemden elde edilen istatistikler aracılığıyla göz önüne alınan evrenin ilgili parametre değerlerinin bulunmasına tahminleme denir. Yani, tahminleme ele alınan evrenden seçilen rassal örneklemlerden elde edilen istatistikler yardımıyla ilgili evrenin dağılımının parametrelerinin belirlenmesi işlemidir. Tahminleme yapabilmek için uyulması gereken işlem adımları aşağıdaki gibidir:

- İlgili evrenden n birimlik rassal örneklem seçilir ve veriler derlenir.
- Bu veriler yardımıyla parametre tahmini için ilgili istatistikler hesaplanır.
- Hesaplanan istatistiğin örnekleme dağılımı yardımıyla parametre değerleri tahmin edilir.

Tahmin bir sayısal değer, tahminleyici ise bir formülasyondur.

İstatistiksel Tahminle Türleri

İstatistiksel tahminleme nokta ve aralık tahmini olarak ikiye ayrılır.

Nokta tahminlemesi, rassal bir örneklemden hesaplanan istatistiğin değerini ilgili evren parametresi $\hat{\theta}$'nın değeri olarak kabul eden tahminlemeye denir. $\hat{\theta}$ ilgili evren parametresinin nokta tahmincisidir. Nokta tahmininde aşağıdaki adımlar izlenir:

- İlgili evrenden n birimlik örneklem rassal olarak seçilir ve veriler derlenir.
- Elde edilen veriler yardımıyla θ parametresinin tahmini için gerekli olan istatistik hesaplanır ve bu istatistiğin değeri θ parametresi için tahmin değeri olarak alınır.

Evren aritmetik ortalaması μ'nün nokta tahminlenmesi, örneklem aritmetik ortalaması olan $\bar{X}$ bir rassal değişkendir. $\bar{X}$'nın değerini bu istatistiğin bilgi ürettiği μ'nün_değerine eşit kabul eden tahminlemeye μ'nün nokta tahmini denir. $\bar{X}$ ise nokta tahmincisi adını alır (S:195, Örnek 1).

Evren oranı Π'nin nokta tahminlenmesi, hakkında bilgi edinmek istenen evren oranı π, örneklem oranı ise p ile gösterilir. Evren oranına ilişkin nokta tahmini, n birimlik bir örneklem için, $p = n/r$ oranının π'ye eşit kabul edilmesidir. Burada ki r binom rassal değişkenidir (S:195, Örnek 2).

Aralık tahminlemesi, aynı evrenden seçilen n birimlik örneklemlerde çok büyük farklılık göstermeyen tahmine güvenilir tahmin denilir. Nokta tahmini sınırlı bir tahminlemedir. Çünkü tahminin güvenirliliği hakkında bilgi veremez. Aralık tahminlemesi, hakkında bilgi edinilmek istenen evren parametresini, belirli bir olasılıkla simetrik bir aralıkta tahmin etme sürecidir.

Örneklem istatistiklerinin ve standart hatalarının değeri seçilen örneklemlere göre değişmektedir. Bu da güven aralığının sınırlarının değişmesine neden olur. Aralık tahmininde, araştırmacı tarafından önceden belirlenen olasılık düzeyine güven düzeyi denir ve 1-α ile gösterilir. α ise anlam düzeyidir. Güven düzeyi güven aralığı tahmininde parametre değerini kapsayan olasılıktır. Aralık tahmini sürecinde izleyen adımlar uygulanır:

- Güven düzeyi belirlenir.
- n birimlik rassal bir örneklem seçilerek, $\hat{\theta}$ hesaplanır.
- $\hat{\theta}$ 'nın dağılımıyla ilgili bilgiler kullanılarak güven aralığı oluşturulur.

Belirlenen bir güven düzeyi için $\hat{\theta}$ istatistiğinin örnekleme dağılımın normal dağıldığı varsayımı altında θ için güven aralığı sınırları;

$\hat{\theta} - z_{\alpha/2}.\sigma_{\hat{\theta}} < \theta < \hat{\theta} + z_{\alpha/2}.\sigma_{\hat{\theta}}$ (σ bilindiğinde)

$\hat{\theta} - z_{\alpha/2}.s_{\hat{\theta}} < \theta < \hat{\theta} + z_{\alpha/2}.s_{\hat{\theta}}$ (σ bilinmediğinde)

formülleri ile bulunur.

Evren aritmetik ortalaması μ'nün aralık tahminlenmesi, evren aritmetik ortalaması için, güven aralığı belirlenmesi işlemine μ'nün aralık tahminlenmesi denir. Bu işlem şu şekildedir.

- Güven düzeyi belirlenir.
- n birimlik rassal örneklem seçilir ve $\bar{X}, s$ istatistikleri hesaplanır.
- $\bar{X}$ 'nın dağılımıyla ilgili bilgiler kullanılarak güven aralığı belirlenir.

Büyük örneklemlerde μ'nün aralık tahminlenmesi, bir örneklemin hacmi n ≥ 30 ise bu örnekleme büyük örneklem denir. Eğer örneklem hacmi yeterince büyük ise evren hangi dağılıma sahip olursa olsun, $\bar{X}$'nın örnekleme dağılımı, ortalaması μ ve varyansı σ^2/n olan normal dağılıma yakınsar. Bu bilgiden hareketle, μ için güven aralığı evren standart sapması bilindiğinde aşağıdaki formül yardımıyla hesaplanabilir.

$$\bar{X} - z_{\alpha/2}.\sigma_{\bar{x}} < \mu < \bar{X} + z_{\alpha/2}.\sigma_{\bar{x}}$$

$\bar{X}, \mu$'nün yansız tahmin edicisidir. Yukarıdaki formüldeki z değeri, araştırmacı tarafından belirlenen güven düzeyine bağlı olarak z tablosundan bulunur (S:198, Tablo 7.1).

$\bar{X}$'nin standart hatası, $\sigma_{\bar{x}}$ örnekleme oranı n/N <0.05 ise ve örnekleme iadeli seçimle gerçekleşmişse ya da ele alınan evren sonsuz ise $\sigma_{\bar{x}} = \sigma/\sqrt{n}$ ile, eğer iadesiz seçim yapılmışsa ve n/N≥ 0.05 ise az önce verilen eşitliğe $\sqrt{N - n/N - 1}$ çarpanı eklenerek hesaplanır.

Eğer evren standart sapması bilinmiyorsa, μ'nün güven aralığı şu formülle hesaplanır.

$$\bar{X} - z_{\alpha/2}.s_{\bar{x}} < \mu < \bar{X} + z_{\alpha/2}.s_{\bar{x}}$$

$s_{\bar{x}}$ 'in tahmini, iadeli seçim olur, n/N<0.05 olur ve ele alınan evren sonsuz olursa $s_{\bar{x}} = s/\sqrt{n}$ ile hesaplanır. Eğer iadesiz seçim yapılır ve n/N$>$ 0.05 olursa yukarıdaki formüle $\sqrt{N-n/N-1}$ eklenerek hesaplanır (S:199, Örnek 3).

Küçük örneklemlerde µ'nün aralık tahminlenmesi, küçük örneklemlerde (n<30), örneklem ortalamasının standart değeri normal dağılmaz. Küçük örneklem durumunda, µ'nün aralık tahmini evrenin normal dağılıp dağılmadığına bağlıdır.

Tanımlanan evren normal dağılıyorsa, değer aralığı $-\infty < (\bar{X}-\mu)/s_{\bar{x}} < +\infty$ olan istatistiği s.d =n-1 serbestlik derecesinde t dağılımı adı verilen sürekli bir dağılıma uyar ve

$$t = \frac{\bar{X}-\mu}{s_{\bar{x}}}$$

şeklinde gösterilir. t dağılımı ortalaması sıfır olan tek modlu ve simetrik bir dağılımdır.

µ için güven aralığı, z istatistiği yerine kullanılan t istatistiği yardımıyla eşitlikteki gibi hesaplanır.

$$\bar{X} - t.s_{\bar{x}} < \mu < \bar{X} + t.s_{\bar{x}}$$

Burada t, 1-α güven düzeyinde n-1 serbestlik dereceli t tablo değeri ve $s_{\bar{x}} = s/\sqrt{n-1}$'dir (S:201, Örnek 4).

Evren oranının aralık tahminlemesi, özetlenen prosedür, büyük örneklem ya da n/N≤ 0.05 olduğu durum içindir. Büyük örneklemler için, p'nin örnekleme dağılımı yaklaşık normal dağılım gösterir ve standart hatası

$$s_p = \sqrt{\frac{p(1-p)}{n}}$$ ile hesaplanır.

π için aralık tahmini şu şekilde yapılır (S:202, Örnek 5):

$$p - z_{\propto/2}.s_p < \Pi < p - z_{\propto/2}.s_p$$

İstatistiksel Hipotez ve İstatistiksel Hipotez Testi

Hipotez bir duruma ilişkin bir önermedir. İstatistiksel hipotez ise, bir dağılımın evren parametresine ilişkin önermedir.

Bir istatistiksel hipotez, doğru ya da yanlış olabilir. Sonucu öğrenebilmek için evren parametresinin bilinmesi (hesaplanması) gerekir. Ancak bu da tam sayımla gerçekleşir. Tam sayım her zaman yapılamadığından örneklemeye başvurulur. Bu durumda istatistiksel hipotezin geçerliliğine karar verebilmek için, bu hipotezlerin ele alınan evrenden seçilen örneklemden hesaplanan $\hat{\theta}$ 'dan ve $\hat{\theta}$ 'nın örnekleme dağılımdan yararlanılarak test edilmesi gerekir.

Bu yapılan işleme istatistiksel hipotez testi denir. Yani, istatistiksel hipotez testi, örneklem istatistiklerinden yararlanarak bir hipotezin geçerliliğini ortaya koymaya yönelik yapılan çalışmalardır.

Örneklemden hesaplanan istatistiğin, parametre hakkında ileri sürülen değere (θ_0) eşit olması beklenemez. Önemli olan $\hat{\theta}-\theta_0$ farkının istatistiksel olarak anlamlı olup olmadığının belirlenmesidir. Eğer anlamlı bir fark belirlenirse hipotez belirli bir hata payıyla reddedilir.

Hipotez Testi Türleri

Hipotez testleri, **parametrik** ve **parametrik olmayan** olarak ikiye ayrılır. *Parametrik testler,* değişkenlerin ölçümünde oranlı ya da eşit aralıklı ölçeğin kullanıldığı testlerdir. En önemli parametrik testler z ve t testleridir. Kitapta sadece parametrik olan testler göz önüne alınmıştır.

Hipotez Testi Sürecinin Adımları

Hipotezlerin ifade edilmesi, *sıfır hipotezi* ve ***karşıt hipotez*** olmak üzere iki çeşit hipotez vardır.

H_0 ile gösterilen sıfır hipotezi, ele alınan evren parametresinin bilinen değerinde herhangi bir değişimin beklenmediğinin ifade edildiği hipotezdir. Sıfır hipotezi $H_0: \theta = \theta_0$ şeklinde ifade edilir.

H_0 'ın test edilebilmesi için H_1 ile gösterilen karşıt hipoteze ihtiyaç vardır. Karşıt hipotez, ele alınan evren parametresinin bilinen değerinde istatistiksel olarak anlamlı bir değişimin beklendiğini ifade eder. Araştırmanın amacına bağlı olarak karşıt hipotez iç şekilde ifade edilebilir. Bunlar; $H_1: \theta \neq \theta_0$, $H_1: \theta < \theta_0$ ve $H_1: \theta > \theta_0$ şeklindedir.

Red bölgesi, sıfır hipotezinin reddedilmesine neden olan örneklem istatistiği ile ilgili değer aralığıdır. Hipotez testleri, karşıt hipotezin ifade ediliş biçimine göre, "iki yönlü", "tek yönlü üst kuyruk" ve "tek yönlü alt kuyruk" testi olarak adlandırılırlar. Bu testlere ait hipotezler aşağıda gösterildiği gibidir:

- İki yönlü testlerde : $H_0: \theta = \theta_0$ ve $H_1: \theta \neq \theta_0$
- Tek yönlü üst kuyruk testinde: $H_1: \theta \neq \theta_0$ ve $H_1: \theta > \theta_0$
- Tek yönlü alt kuyruk testinde: $H_1: \theta > \theta_0$ ve $H_1: \theta < \theta_0$

Anlamlılık düzeyinin belirlenmesi, bir istatistiksel hipotez testinde sıfır hipotezi ya kabul edilir ya da reddedilir. Hipotez testlerinde, sıfır hipotezinin yanlışlıkla kabul ya da reddedilmesi durumuna "çıkarsama hatası" denir. Gerçekte doğru olan sıfır hipotezinin reddedilmesi durumunda I. tip hata (α tipi hata) yapılmış olur. Gerçekte yanlış olan sıfır hipotezinin kabul edilmesi durumunda ise II. tip yani β hatası gerçekleşmiş olur. α tipi hata yapmanın maksimum olasılığına "testin anlamlılık düzeyi" denir.

Anlamlılık düzeyinin belirlenmesi, örneklemden elde edilen bilgiler ışığında doğru olan sıfır hipotezinin reddedilmesi olasılığını belirleyen α'nın seçimidir. Araştırmalarda α genellikle %5 ya da %1 olarak seçilir.

Bu seçim doğru olan H_0 hipotezinin reddedilme olasılığını gösterir.

Verilerin derlenmesi, hipotezler ifade edilip α belirlendikten sonra, evrenden seçilecek örneklemin büyüklüğü kararlaştırılır. Evrenden örneklem seçilerek veriler derlenir ve test istatistiğinin hesabında kullanılır.

Test istatistiğinin seçilmesi, sıfır hipotezinin test edilebilmesi için, örneklem istatistiğinin dağılımının bilinmesi gerekir. Ayrıca uygun bir test istatistiğine de ihtiyaç vardır.

Test istatistiği, örneklem istatistiği $\hat{\theta}$ ile θ_0 arasındaki farkı standart hata birimiyle ifade eden ölçü olarak ifade edilebilir. Örneklemin sıfır hipotezine ne kadar uyduğunun göstergesi olan test istatistiği, test sonunda verilecek istatistiksel kararın dayandığı bir örneklem istatistiğidir.

İstatistiksel kararın verilmesi, istatistiksel karar vermek α anlamlılık düzeyinde sıfır hipotezinin kabul ya da reddedilmesi demektir. Bu kararı verebilmek için gerekli olan ölçüte "test istatistiğinin kritik değeri" adı verilir. Bu ölçüt, bir örnekleme dağılımında ret bölgesin başlama noktasını gösteren değerdir.

Kritik değer seçilen anlamlılık düzeyine, karşıt hipotezin nasıl kurulduğuna ve örneklem istatistiğinin dağılımına göre farklı değerler alabilir. İzleyen açıklamalar $\hat{\theta}$ test istatistiği ve bu istatistiğin standart değeri olan $z = \frac{\hat{\theta}-\theta_0}{\sigma_{\hat{\theta}}}$ istatistiğinin normal dağılıma sahip olduğu ve $\alpha = 0.05$ olarak seçilen durum için yapılmıştır.

Eğer karşıt hipotez $H_1\colon \theta \neq \theta_0$ biçiminde ifade edilmişse, ret bölgesi $\hat{\theta}$ test istatistiğine ilişkin normal dağılımın her iki ucunda simetrik olarak tanımlanmış ve her bölgenin alanı oransal olarak $\alpha/2 = 0{,}025$ olur. Bu durumda kritik değer $z_{tab} = z_{0.5-0.025} = z_{0.475} = 1{,}96$ olur.

İki yönlü testte sıfır hipotezinin reddedilebilmesi için;

$z_{hes} = \left|\frac{\hat{\theta}-\theta_0}{\sigma_{\hat{\theta}}}\right| > z_{tab} = 1{,}96$ olması gerekir. Aksi halde sıfır hipotezi kabul edilir (S:209, Şekil 7.4).

Karşıt hipotez $H_1\colon \theta < \theta_0$ veya $H_1\colon \theta > \theta_0$ şeklinde ifade edilmişse, test istatistiğinin kuyruğunda tanımlanmış alan ilk durum için üst kuyrukta $\alpha = 0.05$, ikinci durum için alt kuyrukta $\alpha = 0.05$ olur. Kritik değer her iki durum içinde simetriktir ve üst kuyrukta pozitif, alt kuyrukta negatiftir. O halde kritik değer $z_{tab} = z_{0.5-0.05} = z_{0.45} = 1{,}64$ olur.

Tek yönlü üst kuyruk testinde eğer,

$z_{hes} = \left(\frac{\hat{\theta}-\theta_0}{\sigma_{\hat{\theta}}}\right) > z_{tab} = 1{,}64$ olursa sıfır hipotezi reddedilir (S:210, Şekil 7.5).

Tek yönlü alt kuyruk testinde ise sıfır hipotezinin reddedilmesi için aşağıdaki durumun gerçekleşmesi gerekir (S:210, Şekil 7.6).

$$z_{hes} = \left(\frac{\hat{\theta}-\theta_0}{\sigma_{\hat{\theta}}}\right) < z_{tab} = -1{,}64$$

Sıfır hipotezinin reddedilmesi, örneklem değeri ile evren parametresi arasında α anlamlılık düzeyinde bir fark olduğu anlamına gelir.

Probleme ilişkin kararın verilmesi, istatistiksel kararın, araştırma problemine ilişkin karara dönüştürülmesi hipotez testlerinde ki önemli noktalardandır.

Tek Evren Parametresiyle İlgili Hipotez Testleri

Bu tür hipotez testlerinde, evrenin göz önüne alınan değişkenine ilişkin bilinen bir parametre değerinin değişmedi şeklindeki sıfır hipotezi test edilir. Örneklem istatistiğinin değeriyle, evren parametresinin bilinen değeri θ_0 karşılaştırılarak karar verilir.

Evren Ortalamasına İlişkin Hipotez Testi, bu testte ilgilenilen evrenden rassal olarak seçilen bir örneklem için hesaplanan ortalama değeriyle ($\bar{X}$), bu örneklemin seçildiği evren aritmetik ortalaması ile ilgili, önceden belirlenen μ_0 gibi bir değer arasındaki farklılığın istatistiksel olarak anlamlı olup olmadığı araştırılır.

Evren Ortalamasına İlişkin Büyük Örneklem Testi, test izleyen adımlarla gerçekleştirilir.

- Rassal olarak bir örneklem seçilir.
- Seçilen örneklem büyük örneklem olmalı ya da göz önüne alınan evrenin normal dağılımlı ve veryansının biliniyor olması gerekir.,
- $H_0\colon \theta = \theta_0$ hipotezi, seçilen anlamlılık düzeyinde test edilir (S:212, Örnek 6, S:214, Örnek 7).

Evren Ortalamasına İlişkin Küçük Örneklem Testi, bazı durumlarda örneklem büyüklüğünü artırmak zor olabilir. Bu durumlarda küçük örneklemler için geliştirilmiş testlere başvurulur. Küçük örneklem olması durumunda σ yerine s'nin kullanılması istatistiksel test üzerinde etkiye sahiptir.

Bu durumda tahmin edilen istatistik $\bar{X} - \mu/s_{\bar{X}}$, n-1 serbestlik derecesiyle t dağılımına sahiptir. t istatistiği ve $s_{\bar{X}}$ aşağıdaki formüllerle hesaplanır.

$$t = \frac{\bar{X}-\mu}{s_{\bar{X}}}, \qquad s_{\bar{X}} = \frac{s}{\sqrt{n-1}}$$

Küçük örneklem kullanılarak yapılan tek evren ortalamasına ait hipotez testlerini, büyük örneklemdeki hipotez testinden ayıran tek fark kullanılan test istatistiğidir (S:216, Örnek 8).

Evren Oranına İlişkin Test, evren oranı π ile gösterilir ve bir evrenin ilgilenilen iki şıklı bir değişkeninin, herhangi bir şıkkına sahip birimlerinin oranına denir. Bu test örneklem oranı p ile evren oranı π'nin iddia edilen değeri π_0 arasındaki farkın istatistiksel olarak anlamlı olup olmadığını araştırır.

Evren oranı π'ye ilişkin testlerde örneklem hacmi büyük ise standartlaştırılmış z istatistiği kullanılır.

$$z = \sqrt{\frac{\mathrm{p} - \Pi_0}{\sigma_p}}$$

Burada σ_p örneklem oranı p'nin örnekleme dağılımının standart sapmasını gösterir ve z izleyen eşitlik yardımıyla hesaplanır (S:218, Örnek 9 ve S:220, Örnek 10).

$$z = \sqrt{\frac{\Pi_0(1 - \Pi_0)}{n}}$$

Giriş

İki değişken arasındaki ilişkiyi tanımlamaya ve ölçmeye **korelasyon analizi** adı verilir. Bir veya birden çok değişkenin başka bir değişken üzerindeki ilişkisini açıklamaya **regresyon analizi** adı verilir.

Regresyon analizinde;

- Bağımlı ve
- Bağımsız değişken olmak üzere iki farklı değişken tipi bulunmaktadır..

Bağımlı değişken, araştırmacının üzerinde çalıştığı değişken olup, bu değişken üzerinde meydana gelen değişimlerin ya da bu değişkenin toplam değişkenliğinin açıklanılmasına çalışılmaktadır.

Bağımsız değişken, ilgilenilen bağımlı değişkende meydana gelen değişim ya da toplam değişkenliğinin üzerinde etkisi olabileceği düşünülen değiken(ler)dir. Regresyon analizinde bir ya da daha çok bağımsız değişken olabilir. Bir tek açıklayıcı değişkenin bulunduğu durum basit doğrusal regresyon ve birden çok açıklayıcı değişkenin bulunduğu durum ise **çoklu doğrusal regresyon** olarak adlandırılır.

İlgilenilen iki değişken arasındaki ilişkinin derecesi için korelasyon analizi kullanılır. Korelasyon analizinin regresyon analizinden farklılık gösterdiği nokta, korelasyon analizinin değişkenler arasındaki ilişkinin yalnızca derecesini göstermesidir. İki değişken arasında yüksek korelasyon olması bu iki değişkenden birinin diğerinin nedeni olabileceğini göstermez.

Korelasyon Analizi

Korelasyon analizi iki değişken arasındaki ilişkinin tanımlanması ve ilişkinin derecesinin belirlenmesidir. İki değişken arasındaki ilişkiyi gözlemlemek için kullanılabilecek en basit yöntemlerden biri bu değişkenlerin dağılım grafiklerini çizmektir (S:228, Şekil 8.1).

İki değişken arasındaki ilişkinin derecesini belirlemek için grafiksel yaklaşım yeterli değildir. Bu amaçla ilişkinin derecesinin belirlenmesi için Pearson korelasyon katsayısı kullanılmaktadır.

Pearson Korelasyon Katsayısı: İki değişken arasındaki doğrusal ilişkinin derecesinin ifade eden korelasyon katsayısı evrendeki x ve y değişkenleri için ρ (ro), örneklemdeki x ve y için ise *r* ile gösterilir. -1 ile +1 arasında değerler alabilen Pearson korelasyon katsayısının, -1 veya +1 değerlerini alması tam ilişkinin varlığını gösterir.

Eksi değerler değişkenler arasında ters yönlü ilişkiyi gösterirken artı değerler aynı yönlü ilişkiyi ifade eder. Pearson korelasyon katsayısının 0 değerini alması ise değişkenler arasında doğrusal ilişkinin olmadığı anlamına gelir.

x ve y değişkenleri arasındaki Pearson korelasyon katsayısı;

$$r = \frac{n\sum_{i=1}^{n} x_i y_i - \left(\sum_{i=1}^{n} x_i\right)\left(\sum_{i=1}^{n} y_i\right)}{\sqrt{n\sum_{i=1}^{n} x_i^2 - \left(\sum_{i=1}^{n} x_i\right)^2}\sqrt{n\sum_{i=1}^{n} y_i^2 - \left(\sum_{i=1}^{n} y_i\right)^2}}$$

eşitliği yardımıyla hesaplanır.

Belirlilik Katsayısı: Bağımlı değişken y'nin bağımsız değişken x tarafından açıklanabilen oranının tespit edilmesi için kullanılır. Evrendeki x ve y değişkenleri için ρ^2 ile gösterilirken, örneklem için r^2 ile gösterilir. Yani örneklem için belirlilik katsayısı, Pearson korelasyon katsayısı olan r 'nin karesinin alınmasıyla hesaplanır. Dolayısıyla belirlilik katsayısı da 0 ile 1 arasında değer alır.

Korelasyon Katsayısı Anlamlılık Testi: Örneklem için hesaplanan korelasyon katsayısı değerinden yararlanarak elde edilen kararın, evren için geçerli olup olmadığı korelasyon katsayısı anlamlılık testi yardımıyla yapılır. Bu test belirtilen adımlar ile gerçekleştirilir.

Adım 1: Hipotezlerin ifade edilmesi

H_0: $\rho = 0$ (x ve y arasında korelasyon yoktur.)

H_1: $\rho \neq 0$ (x ve y arasında korelasyon vardır.)

Adım 2: Anlamlılık düzeyinin belirlenmesi
Doğru olan sıfır hipotezinin (H_0), örneklemden elde edilen bilgiye dayanarak reddedilme olasılığını belirleyen α 'nın seçilmesidir. Bu değer sosyal bilim araştırmalarında genellikle %5 ya da % 1 olarak seçilmektedir.

Adım 3: İstatistiksel test
Örneklemden korelasyon değeri yardımıyla;

$$t = \frac{r\sqrt{n-2}}{1-r^2}$$

istatistiği hesaplanarak, *n-2* serbestlik derecesi ve belirlenen anlam düzeyine göre t dağılımı tablosundan tespit edilen kritik değer ile karşılaştırılır.

Adım 4: İstatistiksel kararın verilmesi
Hesaplanan t'nin mutlak değeri, t tablosu ile belirlenen kritik değerden küçükse H_0 kabul edilir, büyükse H_0 reddedilir, H_1 kabul edilir.

Basit Doğrusal Regresyon Analizi

x ve y değişkenleri kullanılarak elde edilen basit doğrusal regresyon modeli;

$$y_i = \alpha + \beta x_i + \varepsilon$$

şeklinde ifade edilir. Bu modelde;

- y_i : bağımlı değişken *y*'nin i'inci gözlem değerini,

- x_i : bağımsız değişken *x*'in i'inci gözlem değerini,
- ε_i : i'inci gözlem için sıfır ortalamalı ve sabit σ standart sapmalı normal dağılıma sahip olduğu varsayılan rassal hatayı,
- α ve β : tahminlenecek parametre değerlerini gösterir.
  - α, x değişkeni 0 değerini aldığında y'nin alacağı değeri,
  - β, x'teki bir birimlik değişimin y'de meydana getirdiği oransal değişmeyi göstermektedir.

Uygulamada evren değerlerinin tamamına erişim mümkün olmadığından örneklem değerleri kullanılır ve örneklem değerleri kullanıldığında kurulan doğrusal regresyon modeli;

$$y_i = a + bx_i + e_i$$

şeklindedir. Burada *a* ve *b* sırasıyla α ve β'nın tahmin değerlerini göstermektedir.

Kurulan regresyon modelinin tahmin değerlerini elde etmek için en küçük kareler tekniği kullanılmıştır. Bu teknik ile gözlemlenen y_i değerleri ile tahmin edilen $\hat{y}_i$ değerleri arasındaki farkın (hata: $e_i = y_i - \hat{y}_i$) karesi minimize edilerek *a* ve *b* tahmin değerlerine ulaşılır. Bu modelde *b* regresyon doğrusunun eğimini iken *a* ise doğrunun *y* eksenini kestiği noktadır.

En küçük kareler tekniği yardımıyla α ve β parametrelerinin tahmin değerleri aşağıdaki eşitlikler ile hesaplanır:

$$b = \frac{n\sum x_i y_i - \sum x_i \sum y_i}{n\sum x_i^2 - \left(\sum x_i\right)^2}$$

$$a = \bar{y} - b\bar{x}$$

Tahminlenen doğrusal regresyon denklemi;

$$\hat{y}_i = a + bx_i$$

şeklinde yazılır.

Basit doğrusal regresyon analizinde elde edilen sonuçların geçerli olması için aşağıda ifade edilen varsayımların sağlanması gerekir.

- ε , *x*'ten istatistiksel olarak bağımsız olmalıdır.
- ε , normal dağılıma sahip olmalıdır.
- ε değişkeninin aritmetik ortalaması sıfır olmalıdır.
- ε_i ve ε_j birbirinden farklı herhangi iki hata değeri olmak üzere, bu değerler istatistiksel olarak bağımsız olmalıdır.
- ε_i değişkenleri x_i'lerin tüm değerleri için sabit varyansa sahiptir.

Tahminin Standart Hatası: Regresyon analizinde y_i ile y'lerin ortalaması, $\bar{y}$, arasındaki fark toplam değişimi ifade eder. Bu toplam değişim miktarı, açıklanabilen değişim ile açıklanamayan değişimin toplamı olarak da ifade edilebilir. Burada açıklanabilen değişim $\hat{y}_i - \bar{y}$ ifadesi olup regresyon doğrusu tarafından açıklanan kısım iken, açıklanamayan değişim ise $e_i = y_i - \hat{y}_i$ olarak ifade edilir ve bu eşitlik i'inci gözlemin hata terimidir. Genel olarak,

Toplam değişim = Açıklanamayan değişim + Açıklanan değişim;

$$(y_i - \bar{y}) = (y_i - \hat{y}_i) + (\hat{y}_i - \bar{y})$$

şeklinde ifade edilebilir (S:237, Şekil 8.6).

Toplam değişim eşitliğinden yararlanarak sapmaların kareleri toplamı;

$$\sum_{i=1}^{n}(y_i - \bar{y})^2 = \sum_{i=1}^{n}(y_i - \hat{y}_i)^2 + \sum_{i=1}^{n}(\hat{y}_i - \bar{y})^2$$

eşitliği yardımıyla elde edilir. Bu eşitliğin sol tarafı genel kareler toplamı (GKT), sağ tarafındaki ilk toplam ifadesi açıklanamayan değişkenlik yani hata kareler toplamı (HKT) ve sağ tarafındaki ikinci toplam ifadesi ise regresyon doğrusu tarafından açıklanabilen değişkenlik yani regresyon kareler toplamı (RKT) olarak adlandırılır. Toplam değişkenlik kısaca;

$$\text{GKT} = \text{HKT} + \text{RKT}$$

olarak da ifade edilebilir.

Tahminin standart hatası hata kareler toplamının serbestlik derecesi *n-2*'ye bölünmesi ile aşağıdaki gibi elde edilir:

$$s_e = \sqrt{\frac{1}{n-2}\sum_{i=1}^{n}(y_i - \hat{y}_i)^2} = \sqrt{\frac{\sum_{i=1}^{n}(e_i)^2}{n-2}}$$

Hataların normal dağıldığı varsayımından yararlanarak bu değer yorumlanabilir. Eğer aynı veri seti için iki adet regresyon doğrusu elde edilmişse bunlardan standart hatası daha küçük olan tercih edilmelidir.

Örneklem Regresyon Doğrusunun Anlamlılık Testi: y_i değerlerini tahmin ederken modelde kullanılan x bağımsız değişkeninin, bağımlı değişkenin açıklanmasında gerçekten katkısının olup olmadığı örneklem regresyon doğrusunun anlamlılık testi yardımıyla araştırılır. Bu amaçta modelde bağımsız değişkenin katsayısı olan β için hipotez testi aşağıda ifade edildiği gibi çift yönlü olarak kurulur.

H_0: $\beta = 0$ (Regresyon doğrusu anlamlı değildir.)

H_1: $\beta \neq 0$ (Regresyon doğrusu anlamlıdır.)

Test için t testi kullanılır. Bu amaçla β 'nın tahmin değeri olan b regresyon katsayısının standart hatası;

$$s_b = s_e \sqrt{\frac{1}{\sum_{i=1}^{n} \left(x_i - \bar{x}\right)^2}}$$

eşitliği aracılığıyla hesaplanır. Sıfır hipotezinin (H_0) testi için kullanılacak *t* istatistiği;

$$t = \frac{b}{s_b}$$

şeklinde hesaplanır. Elde edilen bu *t* istatistiği n-2 serbestlik derecesi ile belirlenen anlam düzeyine göre tablodan tespit edilen kritik değer ile karşılaştırılır. *t* istatistiği tablo kritik değerinden büyükse sıfır hipotezi (H_0) reddedilir. Yani β 'nın tahmin değeri olan *b* regresyon katsayısı istatistiksel olarak anlamlıdır ve elde edilen regresyon modeli kullanılabilir. *t* istatistiği tablo kritik değerinden küçükse sıfır hipotezi kabul edilir, yani kurulan regresyon modeli kullanılamaz.

Kurulan evren regresyon modelinin eğimi olan β için güven aralığı aşağıdaki şekilde ifade edilir:

$$\left\{ b - t_{\left(\frac{\alpha}{2}, n-2\right)} s_b \leq \beta \leq b + t_{\left(\frac{\alpha}{2}, n-2\right)} s_b \right\} = 1 - \alpha$$

şeklinde ifade edilir.

1. Belirli bir olayın gerçekleşme olasılığının, eldeki mevcut bilgilere dayalı olarak belirlendiği olasılık yaklaşımına aşağıdakilerden hangisi denir?

A) Öznel Olasılık
B) Deneysel Olasılık
C) Nesnel Olasılık
D) Klasik Olasılık
E) Normal Olasılık

Çözüm: *Belirli bir olayın gerçekleşme olasılığının eldeki mevcut bilgilere dayalı olarak belirlendiği olasılık yaklaşımına öznel olasılık denir. Doğru cevap A'dır.*

2. Aşağıdakilerden hangisi kesikli rassal değişkenler için bir olasılık dağılımıdır?

A) Normal Dağılım
B) Gamma Dağılımı
C) Standart Normal dağılım
D) Binom Dağılımı
E) Üstel Dağılım

Çözüm: *Normal dağılım, Gamma dağılımı, Standart normal dağılım, Üstel dağılım 4'ü sürekli rassal değişkenler için olasılık dağılımlarıdır. Buradakilerden sadece binom dağılımı kesikli rassal değişkenler için olasılık dağılımıdır. Doğru cevap D'dir.*

3. A olayının gerçekleşme olasılığı 0,60, B olayının gerçekleşme olasılığı 0,35'dir. A veya B olayının gerçekleşme olasılığı 0,40 olduğuna göre; "A ve B olaylarının birlikte gerçekleşme olasılığı" kaçtır?

A) 0,50
B) 0,55
C) 0,60
D) 0,65
E) 0,40

Çözüm: *P(A U B) = P(A) + P(B) – P(A n B) denkleminden yola çıkarak P(A) = 0,60 ; P(B) = 0,35 P(A U B) = 0,40 ; P(A n B) = P(A) + P(B) – P(A U B) = 0,60+0,35-0,40 = 0,55 olur. Doğru cevap B'dir.*

4. Bir yönetici işçilerinin %80'nine özel sigorta yapıyor. Özel sigorta yaptığı işçilerin %20'sine ek ücret de ödediğine göre, bu fabrikada çalışan işçilerden rassal olarak seçilen bir işçinin hem özel sigortalı olma hem de ek ücret ödenme olasılığı kaçtır?

A) 0,15
B) 0,16
C) 0,25
D) 0,20
E) 0,10

Çözüm: *Genel çarpma kuralına göre A ve B gibi iki olay için, bu olayların birlikte gerçekleşme olasılığı, A olayının gerçekleşme olasılığı ile A'nın gerçekleştiği bilindiğine göre B'nin koşullu olasılığının çarpımına eşittir. A olayı, özel sigorta yapılan işçi olsun. B olayı, bu işçinin ek ücret alma olasılığı olsun. Seçilen işçinin özel sigortasının olma olasılığı P(A) = 0,80 ve seçilen işçinin özel sigortası olduğu bilindiğinde ek ücret alma olasılığı P(B|A) = 0,20 olacaktır. Buna göre; P(A ve B) = P(A). P(B|A) = 0,80*0,20 = 0,16 olarak hesaplanır. Doğru cevap B'dir.*

5. Bir uygulama ödevi için 10 öğrenci arasından bir uygulamacı ve bir teorik araştırmacı kaç farklı şekilde seçilebilir?

A) 80
B) 70
C) 90
D) 45
E) 20

Çözüm: *Uygulama için 10, teorik araştırmacı için 9 kişi söz konusudur. Permütasyon kuralına göre n=10 ve r=2 olmak üzere,* $P_2^{10} = \frac{10!}{(10-2)!}$ *= 10*9 =90 olarak hesaplanır. Doğru cevap C'dir.*

6. "Bir müşteri 10 farklı meyve arasından 3 tane almak istiyor. Müşteri alacağı meyveleri kaç farklı şekilde seçebilir?

A) 100
B) 80
C) 60
D) 110
E) 120

Çözüm: *Seçilen meyvelerin seçim sırası önemli değildir. Yapılabilecek farklı seçim sayısı;*
$C_3^{10} = \binom{10}{3} = \frac{10!}{3!(10-3)!}$ *= 120 olarak elde edilir. Doğru cevap E'dir.*

7. Aşağıdakilerden hangisi sürekli rassal değişkene verilebilecek bir örnektir?
   A) Volkanik dağın sıcaklığı
   B) Bir odadaki insan sayısı
   C) Bir fabrikada aylık üretim sayısı
   D) Bir kişinin sahip olduğu çocuk sayısı
   E) Öğrencilerin aldığı notlar

Çözüm: *Sayılamayacak ya da sonsuz sayıda olası değeri bulunan ve bir sayı aralığı ya da aralık kümesi üzerinde tanımlanan rassal değişkenlere "sürekli rassal değişken" adı verilir. Volkanik dağın sıcaklığı buna bir örnektir. Doğru cevap A'dır.*

8. Bir rassal değişkenin alabileceği değerler;
   2, 8 ve 10; bu değerleri alması olasılıkları sırasıyla 0,3, 0,5 ve 0,2 olduğuna göre, bu olasılık dağılımının aritmetik ortalaması ve varyansı aşağıdakilerden hangisinde doğru olarak verilmiştir?
   A) μ = 6,5 σ = 8,64
   B) μ = 6,6 σ = 9,64
   C) μ = 4,5 σ = 7,50
   D) μ = 8,5 σ = 9,48
   E) μ = 5,5 σ = 8,78

Çözüm: $\mu = \sum_{i=1}^{n} X_i \; Pi(x)$

$\sigma = \sum_{i=1}^{n} (X_i - \mu)\hat{}2 \, Pi(x)$

*μ = (2*0,3)+(8*0,5)+(10*0,2) = 6,6*

$\sigma = [(2-6{,}6)^2 * 0{,}3] + [(8-6{,}6)^2 * 0{,}5] + [(10-6{,}6)^2 * 0{,}2]$ *= 9,64 olarak elde edilir. Doğru cevap B'dir.*

9. Bir binom denemesinde deneme sayısı 6 ve tek bir denemede ilgilenilen sonucun gerçekleşme olasılığı 0,30 ise bu 6 denemede ilgilenilen sonuç sayısının 2 olma olasılığı kaçtır?
   A) 0,25
   B) 0,45
   C) 0,32
   D) 0,20
   E) 0,30

Çözüm: *Binom dağılımında,*

$P(x) = C_x^n \, p^x \, q^{n-x}$ *eşitliği ile olasılık elde edilir.*

$P(x) = \binom{6}{2} (0{,}30)^2 (0{,}70)^{(6-2)}$ *= 0,324 olarak hesaplanır. Doğru cevap C'dir.*

10. Normal dağılıma sahip bir rassal değişkenin aritmetik ortalaması 90 ve varyansı 9 ise bu rassal değişkenin 96'dan fazla olma olasılığı kaçtır?
   A) 0,4772
   B) 0,0228
   C) 0,0128
   D) 0,4775
   E) 0,0356

Çözüm: *Rassal değişkenin 96'dan fazla olma olasılığını bulabilmek için öncelikle bu X değerine karşılık gelen z değeri hesaplanmalıdır. Diğer bir deyişle, z dönüşümü uygulanmalıdır. x=96 noktasına karşılık gelen z değeri,* $z = \frac{X-\mu}{\sigma} = \frac{96-90}{3}$ *= 2 olarak elde edilir. Buna göre soruda istenen olasılık; z=2 değerinin sağında kalan bölgenin alanı olacaktır. Dikkat edilirse aritmetik ortalamanın (ya da z=0 noktasının) sağında ve normal dağılım eğrisinin altında kalan bölgenin alanı, toplam alanın yarısı olan 0,5'tir. Standart normal dağılım tablosuna göre z=0 ile z=2 arasındaki bölgenin alanı 0,4772 olarak bulunur. Buna göre, soruda istenen olasılığı bulabilmek için 0,5 değerinden, z=2 değerine karşılık gelen alanı, yani 0,4772 değerini çıkartmak gerekir. 0,5 – 0,4772 = 0,0228 olarak hesaplanır. Doğru cevap B'dir.*

1. Tam sayım sonucu elde edilen veriler kullanılarak hesaplanan sayısal değerlere ne ad verilir?

A) Örnek istatistiği
B) Parametre
C) Evren
D) Örnekleme
E) Örneklem

Çözüm: *Tam sayım sonucu elde edilen veriler kullanılarak hesaplanan sayısal değerlere parametre adı verilir. Doğru cevap B'dir.*

2. I. Maliyet

II. İncelenecek birimlerin zarara uğraması

III. Zaman

IV. Evreninin tanımlanamaması

Yukarıda verilenlerden hangileri örneklemeyi gerekli kılan nedenlerdendir?

A) Yalnız I
B) I ve III
C) II ve III
D) I, II ve III
E) I, III ve IV

Çözüm: *Maliyet, incelenecek birimlerin zarara uğraması, zaman, doğru bilgi etme ve evreni oluşturan birimlerin değişkenliği örneklemeyi gerekli kılan nedenlerdir. Bu nedenle, doğru cevap D'dir.*

3. I. Evrenin tanımlanması

II. Örneklem hacminin belirlenmesi

III. Çerçevenin belirlenmesi

IV. Örnekleme yönteminin seçilmesi

V. Örneklemenin uygulanması

Örnekleme sürecinin aşamaları aşağıdakilerin hangisinde doğru olarak verilmiştir?

A) I, III, IV, II, V
B) I, II IV, III, V
C) V, III, IV, II, I
D) I, III, V, II, IV
E) II, III, IV, I, V

Çözüm: *Örnekleme yönteminin aşamaları, evrenin tanımlanması, çerçevenin belirlenmesi, örnekleme yönteminin seçilmesi, örneklem hacminin belirlenmesi ve örneklemenin uygulanması şeklindedir. Doğru cevap A'dır.*

4. Örnekleme yöntemleri ile ilgili olarak aşağıda verilen eşleştirmelerden hangisi <u>yanlıştır?</u>

A) Basit rassal örnekleme → Olasılıklı örnekleme
B) Kartopu örneklemesi → Olasılıklı olmayan örnekleme
C) Kolayda örnekleme → Olasılıklı olmayan örnekleme
D) Kota örneklemesi → Olasılıklı olmayan örnekleme
E) Sistematik örnekleme → Olasılıklı olmayan örnekleme

Çözüm: *Basit rassal örnekleme ve sistematik örnekleme olasılıklı örnekleme yöntemlerindendir. Kartopu, kota ve kolayda örnekleme ise olasılıklı olmayan örnekleme yöntemlerindendir. Doğru cevap D'dir.*

5. 6 birimlik bir evrenden 4 birimlik kaç farklı örneklem seçilebilir?

A) 5
B) 10
C) 15
D) 20
E) 25

Çözüm: *10 birimlik bir evrenden 4 birimlik örneklem seçilebilir. Doğru cevap C'dir.*

6. Normal dağılıma sahip bir evrenden 100 birimlik rassal bir örneklem seçilmiştir. Seçilen örneklemin ortalaması 12, standart sapması 2 olarak bulunmuştur. Bu durumda, örneklem $\overline{X}$ nın standart hatası aşağıdakilerden hangisidir?

A) 0.1
B) 0.2
C) 0.3
D) 0.4
E) 0.5

Çözüm: *Örneklem ortalamasının standart hatası* $s_{\overline{X}} = \frac{s}{\sqrt{n}}$ *formülü ile bulunur. Buradan,* $s_{\overline{X}} = \frac{s}{\sqrt{n}} = \frac{2}{\sqrt{100}} = \frac{2}{10} = 0.2$ *elde edilir. Doğru cevap B'dir.*

7. Evren oranını tahmin etmek için kullanılan örnek istatistiği aşağıdakilerden hangisidir?

A) Örneklem ortalaması
B) Örneklem oranı
C) Örnekleme oranı
D) Örneklem varyansı
E) Örneklem standart sapması

Çözüm: *Evren oranını tahmin etmek için örneklem oranı kullanılır. Doğru cevap B'dir.*

8. Örneklem standart sapması hangi parametreyi tahmin etmek için kullanılır?

A) Aritmetik ortalama
B) Oran
C) Aritmetik ortalamalar arasındaki fark
D) Standart sapma
E) Oranlar arasındaki fark

Çözüm: *Örneklem standart sapması, evren standart sapması tahmin etmek için kullanılan örnek istatistiğidir. Doğru cevap D'dir.*

9. Bir otel zinciri, müşterilerin memnuniyet oranını tahmin etmek istemektedir. Bu amaçla, 500 müşteri rassal olarak seçilmiş ve 366 kişi memnun olduğunu bildirmiştir. Bu verileri kullanarak istenen tahminleme yapılırken işlenecek hata düzeyi nedir?

A) 0.7300
B) 0.2100
C) 0.0175
D) 0.0730
E) 0.0210

Çözüm: *p nin standart hatası Tahminleme yaparken işlenebilecek hata düzeyini belirleme imkanı veren istatistik standart hatadır. Dolayısıyla, bu soruda örneklem ortalamasının standart hatasının bulunması gerekmektedir. Örneklem ortalaması p nin standart hatası*

$$s_p = \sqrt{\frac{p(1-p)}{n}}$$

formülü ile hesaplanır. Burada,

$$p = \frac{r}{n} = \frac{366}{500} = 0.73$$

olup $1 - p = 0.27$ dir. Böylece,

$$s_p = \sqrt{\frac{p(1-p)}{n}} = \sqrt{\frac{(0.73).021)}{500}} = 0.0175$$

olarak elde edilir. Doğru cevap C'dir.

10. Ortalaması (μ) 10 olan bir evrenden100 birimlik rassal bir örneklem seçilmiştir. Seçilen örneklemin ortalaması 12, standart sapması 2 olarak bulunmuştur. Bu durumda, örneklem ortalamasının standart z değeri aşağıdakilerden hangisidir?

A) 2
B) 4
C) 6
D) 8
E) 10

Çözüm: *Örneklem ortalamasının standart hatası*

$$s_{\bar{X}} = \frac{s}{\sqrt{n}}$$

formülü ile bulunur. Buradan,

$$s_{\bar{X}} = \frac{s}{\sqrt{n}} = \frac{2}{\sqrt{100}} = \frac{2}{10} = 0.2$$

elde edilir. Ayrıca, soruda verilenlerden $\mu = 10$ dur. Buradan,

$$z = \frac{\bar{X} - \mu}{s_{\bar{X}}} = \frac{12 - 10}{0.2} = 10$$

olur. Doğru cevap E'dir.

ANADOLU ÜNİVERSİTESİ AÇIKÖĞRETİM SİSTEMİ

1. Evren aritmetik ortalaması μ nün nokta tahminleyicisi aşağıdakilerden hangisidir?

A) Örneklem oranı
B) Evren oranı
C) Örneklem aritmetik ortalaması
D) Örneklem varyansı
E) Örneklem standart sapması

Çözüm: *Evren aritmetik ortalaması μ nün nokta tahminleyicisi örneklem aritmetik ortalamasıdır. Doğru cevap C'dir.*

2. I. Örneklemin seçilmesi, verilerin derlenmesi ve test istatistiğinin belirlenmesi

II. Probleme ilişkin kararın verilmesi
III. Hipotezlerin ifade edilmesi
IV. İstatistiksel kararın verilmesi
V. Anlamlılık düzeyinin belirlenmesi

Aşağıdakilerden hangisinde hipotez testinin aşamaları doğru sırada verilmiştir.

A) III, V, I, IV, II
B) IV, I, V, II, III
C) I, II, III, IV, V
D) III, II, I, IV, V
E) V, III, II, I, IV

Çözüm: *Hipotez testinin aşamaları hipotezlerin ifade edilmesi, anlamlılık düzeyinin belirlenmesi, örneklemin seçilmesi, verilerin derlenmesi ve test istatistiğinin belirlenmesi, istatistiksel kararın verilmesi ve probleme ilişkin kararın verilmesi şeklindedir. Bu durumda, doğru sıralama III, V, I, IV ve II'dir. Doğru cevap A'dır.*

3. Bir hipotez testi sürecinde bir tip hata yapması olasılığı 0.10 dur. Bu durumda, anlam düzeyi kaçtır?

A) 0.025
B) 0.005
C) 0.01
D) 0.05
E) 0.10

Çözüm: *Anlam düzeyi, birinci tip yapma olasılığına eşittir. Bu nedenle, anlam düzeyi 0.10'dur. Doğru cevap E'dir.*

4. Bir fabrikada üretilen ürünlerden kusurlu olanların oranının en fazla %1 olduğu iddia edilmektedir.

Bu durumda, hipotezler aşağıdakilerin hangisinde doğru olarak verilmiştir?

A) $H_0: \Pi = 0.01$
$H_1: \Pi < 0.01$
B) $H_0: \Pi = 0.01$
$H_1: \Pi > 0.01$
C) $H_0: \Pi = 0.01$
$H_1: \Pi \neq 0.01$
D) $H_0: \Pi = 0.99$
$H_1: \Pi < 0.99$
E) $H_0: \Pi = 0.99$
$H_1: \Pi = 0.01$

Çözüm: *Bir fabrikada üretilen ürünlerden kusurlu olanların oranının ile ilgilenildiğinden örneklem oranını için hipotez testi yapılması gerekir. Kusurlu vidaların oranı en fazla %1 olduğundan hipotezler* $H_0: \Pi = 0.01$ ve $H_1: \Pi < 0.01$ *şeklinde olmalıdır. Doğru cevap A'dır.*

5.-7. soruları aşağıda verilen bilgiyi kullanarak çözünüz.

Bir hastalıktan normal koşullarda iyileşme süresi ortalama 5 gündür. Yeni geliştirilen bir ilacın iyileşme sürecini azalttığı iddia edilmektedir. Bu amaçla, ilgili hastalığa yeni yakalanmış 15 hasta rassal olarak seçilmiş ve bu hastalara yeni ilaç verilmiştir. Hastaların ortalama iyileşme süreleri 4.7 gün ve standart sapması 0.9 gün olarak bulunmuştur. İlacın iyileşme süresini azaltıp azaltmadığı α=0.05 anlam düzeyinde sınanacaktır.

5. Bu sınamada sıfır hipotezi nedir?

A) μ=5
B) μ>5
C) μ<5
D) μ=4.8
E) μ>4.8

Çözüm: *"Bir hastalıktan normal koşullarda iyileşme süresi ortalama 5 gündür." ifadesinden dolayı sıfır hipotezi μ=5 şeklinde olmalıdır. Doğru cevap A'dır.*

6. Bu sınamada alternatif hipotezi nedir?

A) μ=5
B) μ>5
C) μ<5
D) μ=4.8
E) μ>4.8

Çözüm: *"Yeni geliştirilen bir ilacın iyileşme sürecini azalttığı iddia edilmektedir." ifadesinden dolayı alternatif hipotez μ<5 şeklinde olmalıdır. Doğru cevap C'dir.*

7. Bu sınamada örnekleme dağılımının red bölgesi aşağıdakilerden hangisidir?

A) Sağ uçta %5 lik alan
B) Sağ uçta %2.5 lik alan
C) Sağ uçta %1 lik alana
D) Sol uçta %5 lik alan
E) Sol uçta %2.5 lik alan

Çözüm: *Red bölgesini alternatif hipotez belirler. Burada alternatif hipotez μ<5 şeklinde tek yönlü hipotez olduğundan red bölgesi sol uçta yer alır ve alanı %5'dir Doğru cevap D'dir.*

8. $H_0: \mu = 1000$ ve $H_1: \mu \neq 1000$ olarak veriliyor. Bu hipotez testi için α=0.01 anlam düzeyinde red bölgesi aşağıdakilerden hangisinde doğru verilmiştir?

A) Sağ uçta %1 lik alan ve sol uçta %1 lik alan
B) Sağ uçta %0.05 lik alan ve sol uçta %0.05 lik alan
C) Sağ uçta %5 lik alan ve sol uçta %5 lik alan
D) Sağ uçta %2.5 lik alan ve sol uçta %2.5 lik alan
E) Sağ uçta %0.25 lik alan ve sol uçta %0.25 lik alan

Çözüm: *Red bölgesini alternatif hipotez belirler. Burada alternatif hipotez μ≠1000 şeklinde çift yönlü hipotez olduğundan red bölgesi sağ uçta %0.05 ve sol uçta %0.05 olmak üzere toplam alanı %1 olacak şekilde belirlenir. Doğru cevap B'dir.*

9. $\alpha = 0.05$ olmak üzere $z_{\alpha/2}$ değeri aşağıdakilerden hangisidir?

A) 1.96
B) 2.57
C) 1.64
D) 1.00
E) 0.05

Çözüm: $\alpha = 0.05$ *olmak üzere* $z_{\alpha/2}$ *değeri z tablosu yardımıyla 1.96 olarak elde edilir. Doğru cevap A'dır.*

10. $\alpha = 0.01$ olmak üzere z_{α} değeri aşağıdakilerden hangisidir?

A) 1.96
B) 2.33
C) 1.64
D) 1.28
E) 0.01

Çözüm: $\alpha = 0.01$ *olmak üzere* z_{α} *z tablosu yardımıyla 2.33 olarak elde edilir. Doğru cevap B'dir.*

1. İki değişken arasındaki Pearson korelasyon katsayısı r=0.91 olarak hesaplanmıştır. Bu sonuca göre aşağıdaki ifadelerden hangisi doğrudur?

A) İki değişken arasında korelasyon yoktur.
B) İki değişken arasında negatif yönlü güçlü korelasyon vardır.
C) İki değişken arasında negatif yönlü zayıf korelasyon vardır.
D) İki değişken arasında pozitif yönlü güçlü korelasyon vardır.
E) İki değişken arasında pozitif yönlü zayıf korelasyon vardır.

Çözüm: *Pearson korelasyon katsayısı, iki değişken arasındaki doğrusal ilişkinin hem yönünü hem de derecesini belirler. Bu katsayının +1'e yakın olması iki değişken arasında pozitif yönlü güçlü bir doğrusal ilişkinin yani korelasyonun olduğunu ifade eder. Doğru cevap D'dir.*

2. - 6. Soruları aşağıda verilen bilgiyi kullanarak çözünüz.

$\hat{y} = -2 + 3x$ basit doğrusal regresyon denklemi veriliyor.

2. Regresyon doğrusunun y eksenini kestiği nokta aşağıdakilerden hangisidir?

A) -2
B) 2
C) 3
D) -3
E) 0

Çözüm: *Regresyon doğrusunun y eksenin kestiği nokta denklemdeki sabit terim olan -2'dir. Doğru cevap A'dır.*

3. Regresyon doğrusunun eğimi kaçtır?

A) -2
B) 2
C) 3
D) -3
E) 3

Çözüm: *Regresyon doğrusunun eğimi, denklemdeki x'in katsayısı olan 3'tür. Doğru cevap C'dir.*

4. x=5 için y'nin tahmin edilmiş değeri nedir?

A) 10
B) 11
C) 12
D) 13
E) 14

Çözüm: *x=5 verilen denklemde yerine yazılırsa*

$$\hat{y} = -2 + 3.5 = -2 + 15 = 13$$

elde edilir. Doğru cevap D'dir.

5. Gerçek gözlem değeri y=17 ve x=5 iken artık değeri ne olur?

A) 8
B) 7
C) 6
D) 5
E) 4

Çözüm: *Artık değeri, gerçek y değeri ile tahmin edilmiş y değeri arasındaki farka eşittir. x=5 için tahmin edilmiş y değeri 13 olarak bulunmuştu. Bu durumda,*

$e = y - \hat{y} = 17 - 13 = 4$ *olur. Doğru cevap E'dir.*

6. y=20 için x kaçtır?

A) 7
B) 6
C) 5
D) 4
E) 3

Çözüm: *Verilen denklemde y=20 yazılır ve x çözülürse 6 olarak bulunur. Doğru cevap B'dir.*

7. n=12 ve hata kareler toplamı 221 olan regresyon doğrusunun değişkenliği s_e kaçtır?

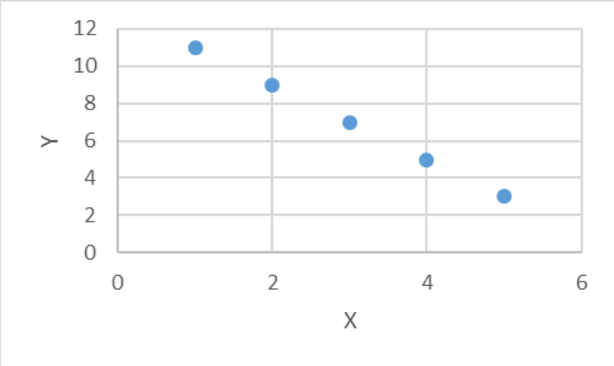
A) 4.29
B) 4.50
C) 4.70
D) 5.12
E) 6.89

Çözüm: *Regresyon doğrusunun değişkenliği s_e aşağıdaki formül yardımıyla hesaplanır.*

$$s_e = \sqrt{\frac{\sum e_i^2}{n-2}} = \sqrt{\frac{221}{12-2}} = \sqrt{\frac{221}{10}} = \sqrt{22.1} = 4.70$$

Doğru cevap C'dir.

8. X ve Y değişkenlerinin serpilme grafiği aşağıda verilmiştir.



Bu durumda, X ve Y arasındaki korelasyon katsayısının aşağıdakilerden hangisi olması beklenir?

A) +0.90
B) +0.50
C) 0
D) -0.30
E) -0.95

Çözüm: *Korelasyon katsayısı, iki değişken arasındaki doğrusal ilişkinin hem yönünü hem de derecesini belirler. Bu katsayının -1'e yakın olması iki değişken arasında negatif yönlü güçlü bir doğrusal ilişkinin yani korelasyonun olduğunu ifade eder. Verilen grafikte X artarken Y azalmaktadır. Dolayısıyla burada negatif yönlü bir ilişki söz konusudur. İlişki doğrusal olduğundan korelasyon katsayısı -1'e yakın bir değer olmalıdır. Doğru cevap E'dir.*

9. İki değişken arasındaki belirlilik katsayısı 0.81 olarak bulunmuştur. Bu durumda, korelasyon katsayısı kaçtır?

A) 0.90
B) 0.65
C) 0.50
D) 0.35
E) 0.10

Çözüm: *Belirlilik katsayısı, korelasyon katsayısının karesine eşittir. Bu durumda, korelasyon katsayısı, belirlilik katsayısının kareköküne eşit olur. 0.81'in karekökü 0.90 olduğundan korelasyon katsayısı 0.90 olarak bulunur. Doğru cevap A'dır.*

10. Evren parametresi β nın değerini tahmin etmek için kullanılan b regresyon katsayısının değeri 2.5 ve standart hatası s_b =0.2 olarak bulunmuştur. Bu durumda, sıfır hipotezini test etmek için kullanılan t istatistiğinin değeri aşağıdakilerden hangisidir?

A) 2.5
B) 12.5
C) 22.5
D) 32.5
E) 35

Çözüm: *Evren parametresi β nın anlamlılığını test etmek için aşağıdaki formül kullanılır.*

$$t = \frac{b}{s_b} = \frac{2.5}{0.2} = 12.5$$ *Doğru cevap B'dir.*

İST203U-İSTATİSTİK

Ünite 5: Olasılık Kuramı

OLASILIK KURAMI

1. *Olasılık nedir?*

Cevap: Sıfır ile bir kapalı aralığında (0 ve 1 dahil) değerler alan ve bir olayın ortaya çıkma şansını belirten sayıya ***olasılık*** adı verilir.

2. *Olasılık değeri nasıl yorumlanmaktadır?*

Cevap: Olasılık değeri sıfıra yaklaştıkça, olayın gerçekleşme şansı gittikçe azalıyor, olasılık değeri bire yaklaştıkça da o olayın gerçekleşme şansı gittikçe artıyor şeklinde yorumlanır.

3. *Deneme, sonuç, örneklem uzayı ve olay kavramlarını açıklayınız.*

Cevap: Çeşitli olası gözlemlerden yalnızca birinin gerçekleşmesi ile sonuçlanan sürece ***deneme*** adı verilir. Bir denemenin iki ya da daha fazla olası sonucu bulunur ve bu sonuçlardan hangisinin gerçekleşeceği belirsizdir. Bir denemenin sona erme biçimine ***sonuç*** adı verilir. Bir denemenin tüm olası sonuçlarından oluşan kümeye ***örneklem uzayı*** adı verilir ve bu küme S harfi ile ifade edilir. Bir denemenin bir ya da daha fazla sonucundan oluşan kümeye ise ***olay*** adı verilir.

OLASILIK TANIMLARI

4. *Olasılık tanımlarına ilişkin yaklaşımlar nelerdir?*

Cevap: Olasılık tanımlara ilişkin nesnel ve öznel olasılık olmak üzere iki farklı yaklaşım bulunmaktadır.

5. *Kaç farklı olasılık tanımı bulunmaktadır ve bu tanımlar nelerdir?*

Cevap: Üç farklı olasılık tanımı bulunmaktadır:

- Klasik olasılık,
- Deneysel Olasılık,
- Özel (Sübjektif) Olasılık.

6. *Klasik Olasılık tanımına göre bir olayın gerçekleşme olasılığı nasıl hesaplanmaktadır?*

Cevap: Klasik olasılık çerçevesinde bir olayın gerçekleşme olasılığı ilgilenilen sonuçların sayısının, olası tüm sonuçların sayısına bölünmesi yoluyla hesaplanır. Klasik olasılık yaklaşımıyla bir A olayının gerçekleşme olasılığı aşağıdaki eşitlik yardımıyla hesaplanabilir:

$$\mathrm{P(A)} = \frac{\mathrm{n}}{\mathrm{N}}$$

Eşitlikte;

- İlgilenilen sonuç sayısı **n** ve
- Olası tüm sonuçların sayısı **N** ile belirtilmiştir.

7. *Bir zarın atılması denemesinde, zarın yüzünün çift gelmesi olasılığı nedir?*

Cevap: Bu deneme için olası sonuçlar: 1, 2, 3, 4, 5, 6 olarak belirlenir.

Bu 6 eşit olasılıklı sonuç içerisinde ilgilenilen sonuçlar "2", "4" ve "6" olup bunların sayısı 3'dir.

A olayı; "Zarın yüzünün çift gelmesi" olarak tanımlandığında;

$$P(A) = \frac{\cancel{3}}{\cancel{6}} = \frac{1}{2} \text{ olur.}$$

8. *Karşılıklı ayrık olaylar nelerdir?*

Cevap: Bir olay gerçekleştiği anda diğer olayların hiçbiri gerçekleşmiyorsa, bu olaylara ***karşılıklı ayrık olaylar*** adı verilir.

9. *Bütüne tamamlayan olaylar nelerdir?*

Cevap: Eğer bir deneme, olası tüm sonuçları içeren bir olaylar kümesinden oluşuyorsa, bu olaylar kümesine ***bütüne tamamlayan olaylar*** adı verilir.

10. *Deneysel olasılık tanımına göre olasılık hesabı nasıl yapılmaktadır?*

Cevap: Deneysel olasılık yaklaşımında bir olayın gerçekleşme olasılığını hesaplamak için, geçmişte benzer olayların gerçekleşme sayısının oranına bakılır. Geçmişte ortaya çıkan olayların sayısı f ile ve toplam gözlem sayısı N ile belirtildiğinde, deneysel olasılık yaklaşımıyla bir A olayının gerçekleşme olasılığı aşağıdaki eşitlik yardımıyla hesaplanabilir:

$$P(A) = \frac{f}{N}$$

11. *Öznel olasılık kavramını açıklayınız.*

Cevap: Öznel Olasılık, belirli bir olayın gerçekleşme olasılığının, eldeki mevcut bilgilere dayalı olarak belirlendiği olasılık yaklaşımıdır.

OLASILIK HESAPLAMA KURALLARI

12. *Toplama kuralları nelerdir ve hangi amaçla kullanılmaktadır?*

Cevap: Toplama Kuralları, olayların birleşimlerinin olasılığını bulmada kullanılır. Bu kurallar aşağıda verilmiştir:

- Özel Toplama Kuralı
- Tümleyen Kuralı
- Genel Toplama Kuralı

13. *Özel Toplama kuralı hangi tür olaylar için kullanılmaktadır ve nasıl ifade edilmektedir?*

Cevap: Özel Toplama Kuralı yalnızca karşılıklı ayrık olaylar için uygulanabilir. Eğer A ve B olayları karşılıklı ayrık olaylar ise, özel toplama kuralına göre bu olaylardan birinin veya diğerinin gerçekleşme olasılığı, bu olayların ayrı ayrı gerçekleşme olasılıklarının toplamına eşittir. Özel toplama kuralına ilişkin eşitlik aşağıda verilmiştir:

$$P(A \text{ veya B}) = P(A) + P(B)$$

14. *Tümleyen kuralına ilişkin matematiksel gösterimi nasıl yapılmaktadır?*

Cevap: A olayının gerçekleşme olasılığı $P(A)$, gerçekleşmeme olasılığı ise $P(\bar{A})$ ile gösterilirse, bu iki olasılığın toplamının 1 olacağı kolaylıkla tahmin edilebilir.

Buna göre, $P(A)+P(\bar{A})=1$ olur.

Bu eşitlikten, $P(A)=1-P(\bar{A})$ olduğu sonucuna varılır. Bu son yazılan eşitlik tümleyen kuralının matematiksel gösterimidir.

15. *Ortak olasılık kavramını açıklayınız.*
Cevap: İki olayın birlikte gerçekleşme olasılığına ***ortak olasılık*** adı verilir. A ve B gibi iki olay için $P(\text{A ve B})$ şeklinde gösterilir.

16. *Genel toplama kuralını ifade ediniz.*
Cevap: A ve B gibi iki olay için genel toplama kuralına ilişkin eşitlik aşağıda verilmiştir:

$$P(\text{A veya B}) = P(A)+P(B)-P(\text{A ve B})$$

$P(\text{A veya B})$ İfadesindeki "veya" kelimesi A'nın ya da B'nin gerçekleşebileceğini öne sürmektedir. Bu ifade, A ve B olaylarının birlikte gerçekleşmesi olan $P(\text{A ve B})$ olasılığını da içermektedir.

17. *Hangi durumda genel toplam kuralı yerine özel toplama kuralı uygulanabilir?*
Cevap: Eğer olaylar karşılıklı ayrık ise, $P(\text{A ve B})$ ortak olasılığı sıfır olur ve bu durumda A veya B'nin gerçekleşme olasılığının hesaplanmasında genel toplama kuralı özel toplama kuralına dönüşecektir.

18. *Çarpma kurallarından ne zaman yararlanılır?*
Cevap: A ve B gibi iki olayın birlikte gerçekleşme olasılığı hesaplanmak istendiğinde, bu iki olayın kesişimine ilişkin olasılık değerlendirilir ve bu olasılığın hesaplanmasında çarpma kurallarından yararlanılır.

19. *Bağımsız ve bağımlı olaylar kavramlarını açıklayınız.*
Cevap: A olayı B olayının gerçekleşme olasılığını etkilemiyorsa, bu A ve B olayları bağımsız olaylardır. Eğer iki olay bağımsız değilse, bu olaylara ***bağımlı olaylar*** adı verilir.

20. *Özel çarpma kuralını ifade ediniz.*
Cevap: Özel Çarpma Kuralına göre; A ve B bağımsız olayları için A ve B'nin birlikte gerçekleşme olasılığı, bu iki olayın ayrı ayrı gerçekleşme olasılıklarının çarpımına eşittir:

$$P(\text{A ve B}) = P(A)\cdot P(B)$$

Özel çarpma kuralı ikiden fazla olay için genelleştirilebilir. Örneğin, A, B ve C bağımsız olayları için özel çarpma kuralı, bu üç olayın birlikte gerçekleşme olasılığını hesaplamada kullanılabilir:

$$P(\text{A ve B ve C}) = P(A)\cdot P(B)\cdot P(C)$$

21. *Koşullu olasılık kavramını tanımlayınız.*
Cevap: B olayının gerçekleştiği bilindiğinde, A olayının gerçekleşme olasılığına ***koşullu olasılık*** adı verilir ve bu olasılık $P(A|B)$ ile gösterilir. $P(A|B)$ ifadesi, *"B bilindiğine göre, A'nın koşullu olasılığı"* olarak okunur.

22. *Genel çarpma kuralını ifade ediniz.*
Cevap: A ve B gibi iki olay için, bu olayların birlikte gerçekleşme olasılığı, A olayının gerçekleşme olasılığı ile A'nın gerçekleştiği bilindiğine göre B'nin koşullu olasılığının çarpımına eşittir:

$$P(\text{A ve B}) = P(A)\cdot P(B|A)$$

BAYES TEOREMİ

23. *Bayes teoremini ifade ediniz.*
A_1 ve A_2 olayları, karşılıklı ayrık ve bütüne tamamlayan olaylar olsun. Bayes teoremine göre, B olayının gerçekleştiği bilindiğine göre, A_i 'nin gerçekleşme olasılığı aşağıdaki eşitlik ile hesaplanır:

$$P(A_i|B) = \frac{P(A_i)P(B|A_i)}{P(A_1)P(B|A_1)+P(A_2)P(B|A_2)} \quad , \quad i=1,2 \text{ için}$$

Bayes teoremi $A_1, A_2, \ldots, A_n$, n adet karşılıklı ayrık ve bütüne tamamlayan olay için genellenirse aşağıdaki eşitlik ile hesaplanır:

$$P(A_i|B) = \frac{P(A_i)P(B|A_i)}{P(A_1)P(B|A_1)+P(A_2)P(B|A_2)+\ldots+P(A_n)P(B|A_n)}$$

SAYMA KURALLARI

24. *Saymanın temel ilkesini açıklayınız.*
Cevap: Eğer bir işlem m farklı yolla, bir başka işlem de n farklı yolla gerçekleşebiliyorsa, bu iki işlem birlikte $m\times n$ farklı yolla gerçekleşir.

25. *Faktöriyel kavramını açıklayınız?*
Cevap: n pozitif tamsayısından küçük ve eşit bütün pozitif tamsayıların çarpımı, **n *faktöriyel*** olarak adlandırılır ve bu ifade "n!" ile gösterilir. Bu tanıma göre; $n! = n(n-1)(n-2)\ldots 2.1 = n(n-1)!$ olarak yazılabilir.

25. *Özel faktöriyel değerleri nelerdir?*
Cevap: Özel olarak $0!=1$ ve $1!=1$ 'dir.

26. *Permütasyon nedir ve nasıl hesaplanır?*
Cevap: n olası nesnenin tek bir grubundan seçilen r adet nesnenin herhangi bir sıralamasına ***permütasyon*** adı verilir.

Toplam farklı permütasyon sayısını elde etmek için gerekli eşitlik aşağıda verilmiştir:

$${}_nP_r = \frac{n!}{(n-r)!}$$

Burada;

- n: toplam nesne sayısını,
- r: seçilen nesne sayısını belirtmektedir.

27. *Kombinasyon nedir ve nasıl hesaplanır?*
Cevap: Seçilecek nesnelerin diziliş sırası önemli değilse, yapılan her bir seçime bir ***kombinasyon*** adı verilir.

Kombinasyon kuralına göre, n adet nesneden oluşan bir kümeden r nesnelik kombinasyon sayısı aşağıdaki eşitlikle hesaplanır:

$$_{n}C_{r} = \frac{n!}{r!(n-r)!}$$

28. *Kombinasyon ve permütasyon arasındaki fark nedir?*
Cevap: Seçilecek nesnelerin diziliş sırası önemli değilse kombinasyon, nesnelerin dizildiği sıra önemli ise permütasyon kullanılmaktadır.

RASSAL DEĞİŞKENLER ve OLASILIK DAĞILIMLARI

29. *Rassal değişkeni tanımlayınız.*
Cevap: Denemeden denemeye farklı değerler alan ve aldığı bu değerleri belli bir olasılıkla alan değişkenlere ***rassal değişken*** adı verilir.

30. *Rassal değişkenlerin ortaya çıkış şekilleri nelerdir?*
Cevap: Rassal değişkenler;

- Kesikli ve
- Sürekli rassal değişkenler olmak üzere iki şekilde ortaya çıkarlar:

31. *Hangi rassal değişkenler kesikli rassal değişken olarak adlandırılır?*
Cevap: Sonlu ya da sayılabilir sayıda farklı değeri bulunan rassal değişkenlere ***kesikli rassal değişken*** adı verilir.

32. *Hangi rassal değişkenler sürekli rassal değişken olarak adlandırılır?*
Cevap: Sayılamayacak ya da sonsuz sayıda olası değeri bulunan ve bir sayı aralığı ya da aralık kümesi üzerinde tanımlanan rassal değişkenlere ***sürekli rassal değişken*** adı verilir.

33. *Olasılık dağılımını tanımlayınız?*
Cevap: Bir denemedeki olası tüm sonuçların ve bu sonuçların her birine ilişkin olasılıkların yer aldığı listeye ***olasılık dağılımı*** adı verilir.

34. *Olasılık dağılımlarının temel özellikleri nelerdir?*
Cevap: İlgili temel özellikler şöyle sıralanabilir:

- Belli bir sonucun olasılığı 0 ile 1 kapalı aralığında değerler alır.
- Tüm karşılıklı ayrık olayların olasılıkları toplamı 1'e eşittir.

35. *Kesikli olasılık dağılımları ve sürekli olasılık dağılımları nasıl oluşmaktadır?*
Cevap: Kesikli bir rassal değişkenin olası değerler kümesi düzenlendiğinde, bu dağılım kesikli olasılık dağılımı olacaktır. Eğer rassal değişken sürekli ise, olasılık dağılımı sürekli olasılık dağılımı adını alır.

36. *Binom dağılımının özellikleri nelerdir?*
Cevap: İlgili özellikler şöyle sıralanabilir:

1. Denemeler, daima aynı koşullarda tekrarlanmalıdır.
2. Yapılacak her denemenin sonunda, var olan karşılıklı ayrık iki sonuçtan yalnızca birisi ortaya çıkmalıdır. Bu sonuçlardan birisi ilgilenilen sonuç, diğeri ise bunun tümleyeni olan ilgilenilmeyen sonuçtur.
3. Rassal değişken, sabit sayıda denemedeki ilgilenilen durumun sayısını belirtir.
4. Tek bir denemede ilgilenilen sonucun gerçekleşme olasılığı, tüm denemelerde aynı kalmalıdır.
5. Denemeler birbirinden bağımsız yapılmalıdır.

37. *Binom dağılımında sıklıkla karşılaşılan olası sonuçlara örnek veriniz?*
Cevap: Binom denemelerindeki olası iki sonuç genellikle;

- "Başarılı-başarısız",
- "Var-yok",
- "Ölü-sağ",
- "Pozitif negatif" vb. ikililerdir.

38. *Sürekli olasılık dağılımlarına örnek veriniz.*
Cevap: Sürekli olasılık dağılımları, bir nesnenin uzunluğu, ağırlığı, sıcaklığı gibi aldığı değerleri genellikle bir ölçüm sonucunda alan sürekli rassal değişkenlere ilişkin dağılımlardır.

39. *Beklenen değer nedir?*
Cevap: Bir olasılık dağılımının ortalamasına, *o olasılık dağılımının **beklenen değeri*** adı verilir.

40. *Kesikli dağılımlar için beklenen değer nasıl hesaplanır?*
Cevap: Kesikli olasılık dağılımları için ortalama aşağıdaki eşitlik yardımıyla hesaplanır:

$$\mu = \sum xP(x)$$

Burada $P(x)$; belli bir x değerine karşılık gelen olasılık değeridir.

41. *Binom olasılık dağılımı nasıl formüle edilir?*
Cevap: Binom olasılık dağılımı aşağıdaki eşitlik ile verilir:

$$P(x) = {}_{n}C_{x}\, p^{x}(1-p)^{n-x}$$

Burada;

- n : Deneme sayısı,
- x : n denemede ilgilenilen sonuç sayısı olarak tanımlanan rassal değişken,
- p : Tek bir denemede ilgilenilen sonucun gerçekleşme olasılığı ve
- $1-p$: Tek bir denemede ilgilenilmeyen sonucun gerçekleşme olasılığıdır.

42. *Normal olasılık dağılımı nasıl formüle edilir?*
Cevap: Normal olasılık dağılımı aşağıdaki eşitlik ile verilir:

$$P(x) = \frac{1}{\sigma\sqrt{2\pi}} e^{-\frac{1}{2}\left(\frac{X-\mu}{\sigma}\right)^2}$$

Burada;

- x : sürekli rassal değişkenin aldığı değeri,
- μ : evren ortalamasını,
- σ : evren standart sapmasını,
- π : pi sayısı olan 3,14'ü ve
- e : doğal logaritma sisteminin tabanı olan 2,718 sayısını belirtmektedir.

Dolayısıyla normal dağılımda, birincisi ortalama, ikincisi de standart sapma olmak üzere iki adet bilinmeyen parametre bulunur.

43. *Sürekli ve kesikli rassal değişkenlerin olasılıklarının hesaplanmasındaki fark nedir?*
Cevap: Sürekli rassal değişkenler, belli bir aralıkta sonsuz sayıda değer alırlar. Dolayısıyla, sürekli rassal değişkenlerde kesikli değişkenler için belli bir değeri alma olasılığı yerine, belli bir aralıkta yer alma olasılığı hesaplanır.

44. *Normal dağılımda mod, medyan ve aritmetik ortalama değerlerinin konumları nasıldır?*
Cevap: Normal dağılımın aritmetik ortalama, mod ve medyan değerleri eşittir ve bu değerler dağılımın merkezinde yer alırlar.

45. *Standart normal dağılım hangi amaçla kullanılmaktadır?*
Cevap: Aritmetik ortalama ve standart sapma değerlerine ilişkin olarak sonsuz sayıda farklı normal dağılım eğrisi çizilebileceğinden, tüm eğriler için olasılık hesabında kullanılacak tabloların oluşturulması mümkün değildir. Bu nedenle, olasılıkların belirlenmesinde standart normal dağılımdan yararlanılır. Normal dağılımın özel bir durumu olan standart normal dağılım, elde edilebilecek tüm normal dağılımlar için olasılık belirlemede kullanılabilmektedir.

46. *Normal dağılım ile standart normal dağılım arasındaki ilişki nedir?*
Cevap: Herhangi bir normal dağılım, gözlem değerlerinden aritmetik ortalama değeri çıkartıldıktan sonra, bu farkın standart sapmaya bölünmesi yoluyla standart normal dağılıma dönüştürülebilir. Aritmetik ortalaması 0 ve standart sapması 1 olan normal dağılıma ***standart normal dağılım*** adı verilir.

47 *z değerleri nedir ve nasıl elde edilir?*
Cevap: Herhangi bir normal dağılım, gözlem değerlerinden aritmetik ortalama değeri çıkartıldıktan sonra, bu farkın standart sapmaya bölünmesi yoluyla standart normal dağılıma dönüştürülebilir. Bu işlem sonucu elde edilen değerlere ***z değerleri*** ya da ***standart normal değerler*** adı verilir.

X gözlem değerleri aşağıdaki eşitlik yardımıyla z değerine dönüştürülür:

$$z = \frac{X-\mu}{\sigma}$$

Burada,

- X : Herhangi bir gözlem ya da ölçüm değerini,
- μ : dağılımın aritmetik ortalamasını,
- σ : dağılımın standart sapmasını belirtir.

48. *Z değerlerinin özellikleri nelerdir?*
Cevap: z değerleri standart değerler olduğu için birimleri yoktur. z değerleri, X gözlem değerlerinin aritmetik ortalamadan kaç standart sapma uzaklıkta olduğunu belirtir.

49. *Normal dağılımın konumu nasıl belirlenmektedir?*
Cevap: Normal dağılımın konumu, μ aritmetik ortalamasına, yayılması ise σ standart sapmasına göre belirlenir.

İST203U-İSTATİSTİK

Ünite 6: Örnekleme ve Bazı Örnekleme Dağılımları

TAM SAYIM VE ÖRNEKLEME

1. *Örnekleme kavramında evren neyi ifade eder?*
Cevap: Hakkında araştırma yapılacak birimler topluluğuna ***evren*** denir.

2. *Tam sayım işlemi nedir ve hangi durumda tam sayım sonucu elde edilen bilgiler kesin ve doğrudur?*
Cevap: Planlanan bir istatistiksel araştırma için tanımlanan sonlu evrenin bütün birimleri üzerinde araştırmaya konu olan değişkenler itibariyle veri derleniyorsa yapılan işleme ***tam sayım*** denir. Tam sayım sonucunda elde edilen veri kümesinin çözümlenmesiyle elde edilen bilgiler veri derleme ve çözümleme hatası işlenmemiş ise kesin ve doğru bilgilerdir.

3. *Örnekleme istatistiği nedir ve ne amaçla kullanılır?*
Cevap: Evreni temsil eden, onun bir modeli olan örneklemden elde edilen veri kümesi kullanılarak yapılacak çözümlemenin sonucu olan bilgi örneklem istatistiğidir. Örneklem istatistiğinin değeri bilinmeyen evren parametresi hakkında genelleme yapmak amacıyla kullanılır.

4. *İyi bir örneklem nasıl ifade edilir, örneklemin birim sayısı küçük olması önemli bir sorun mudur?*
Cevap: İyi bir örneklem, eğer evren doğru ve net tanımlanmış ise örneklemin araştırmaya konu olan değişkenler itibariyle evreni iyi temsil etmesidir. Örneklem birim sayısı küçük olması durumunda bile evreni iyi temsil ediyorsa bu bir sorun olarak nitelendirilemez.

ÖRNEKLEMEYİ GEREKLİ KILAN NEDENLER

5. *Tamsayım yerine örneklemeyi gerekli kılan nedenler nelerdir?*
Cevap: Örneklemeyi gerekli kılan nedenler şöyle sıralanabilir:

- Evrenin sonsuz olması,
- Maliyet,
- Zaman.

6. *Maliyet bakımından tam sayım yerine örneklemenin tercih edilmesinin nedeni nedir?*
Cevap: Örnekleme bütçesinin kısıtlı olması örneklemeyi tam sayıma tercih etmede en önemli belirleyicilerdendir.
Örnekleme tam sayıma göre daha az maliyetle bilgi üretme imkanı sağlar. Öte yandan eğer evren hacmi küçükse veya tam sayım yapmak bütçe olanaklarıyla da mümkünse tam sayım tercih edilmelidir. Burada dikkat edilmesi gereken nokta tam sayım yapma maliyetinin, elde edilecek bilginin değerinden küçük olması gerekir. Aksi durumda örneklemeye başvurmak uygun olacaktır.

7. *Bilgiye çok hızlı gereksinim olduğunda örnekleme mi tam sayım mı kullanılır?*
Cevap: Örnekleme, tam sayıma göre daha kısa zamanda ve yeterli ayrıntıda bilgi elde etme olanağı verir. Örnekleme bilgiye çok hızlı gereksinim olduğu durumlarda özellikle önemlidir.

8. *Evreni oluşturan birimler değişken olduğunda örneklemede nasıl bir yol izlenmelidir?*
Cevap: Evreni oluşturan birimler araştırmaya konu olan değişkenler bakımından heterojen olduğunda yani değişkenler arasındaki farklılık yüksek olduğunda mümkün ise tam sayım yapmak, değil ise büyük hacimli örneklem seçmek gerekir.

9. *Keyfi örnekleme nedir, bir örnek ile açıklayınız?*
Cevap: Örneklem oluşturulurken tanımlanan evreni oluşturan birimler arasında fark gözetilir, yani bütün birimlere bilinen bir olasılıkla seçilme şansı verilmez ise bu türden birim seçimine ***keyfi örnekleme*** adı verilir.

Bu seçim yönteminde araştırmacı, hangi birimlerin örnekleme seçileceğini bilerek ve isteyerek belirler. Örneğin bir üniversitedeki öğrencilerin sorunlarını belirlemek amacıyla yapılacak bir araştırma için her bir bölümde yalnızca tanıdığınız öğrencileri seçme yoluyla bir örneklem oluşturursanız yapmış olduğunuz seçim keyfi örneklemeye bir örnektir.

10. *Sonlu evrenlerde rassal örneklem seçim yöntemleri nelerdir?*
Cevap: Rassal örneklem seçim yöntemleri şöyle sıralanabilir:

- Kura seçimi,
- Sistematik seçim.

11. *Olasılıklı örneklemenin üstünlükleri nelerdir?*
Cevap: Olasılıklı örneklemenin üstünlükleri şöyle sıralanabilir:

- Örneklemden elde edilen verilerden hesaplanan istatistikler evren parametreleri hakkında genelleme yapmak üzere kullanılabilir.
- Örneklem hatasının büyüklüğü hakkında bilgi elde edilebilir.
- Keyfi seçimde söz konusu olabilecek yanlılık giderilmiş olur.

12. *Etkinlik ve doğruluk kavramlarını açıklayınız?*
Cevap: Etkinlik, örnekleme maliyeti ve doğruluğu arasındaki dengeyi yansıtan bir kavramdır. Doğruluk ise ölçülecek özelliğin belirsizliği ile ilgili düzeyi gösterir.

Doğruluk ile örnekleme hataları arasında ters ilişki varken maliyetle aynı yönde ilişki vardır. Yani daha çok maliyet daha doğru bilgi, daha doğru bilgi daha az hatalı karar ve tahmin demektir.

ÖRNEKLEME SÜRECİNİN AŞAMALARI

13. *Örnekleme sürecinin aşamaları nelerdir?*
Cevap: Örnekleme sürecinin aşamaları şöyle sıralanabilir:

- Evrenin tanımlanması,
- Çerçevenin belirlenmesi,
- Örnekleme yönteminin seçilmesi,
- Örneklem hacminin belirlenmesi,
- Örneklemenin uygulanması.

14. *Evrenin tanımlanması aşamaları nelerdir?*
Cevap: Evrenin tanımlanması aşamaları şöyle sıralanabilir:
- Örnekleme birimi,
- Gözlem birimi,
- Yer ve zaman kavramı.

15. *Örneklemede evren kavramı nasıl açıklanır?*
Cevap: Araştırmacı tarafından belirlenen bir tanıma uyan ve hakkında bilgilerin üretileceği, çıkarımların yapılacağı birimlerden oluşan topluluktur. Evrenin ayrıntılı bir biçimde tanımlanmasıyla, hangi birimlerin araştırma kapsamına alınacağı belirlenmiş olur.

16. *Gözlem birimi ve örnekleme birimi ayrımına gidilmesinin nedeni nedir?*
Cevap: Gözlem birimi ve örnekleme birimi ayrımına gidilmesinin nedeni gözlem birimleri ile ilgili bir çerçevenin temine edilmesinin veya hazırlanmasının zor, maliyetli ve çok zaman alacak olmasıdır.

Örnekleme birimi birden çok gözlem birimini kapsayacak şekilde de tanımlanabilir. Örneğin, öğrencilerin beslenme programlarıyla ilgili yapılan bir araştırmada beslenme programıyla ilgili öğrencinin hem annesinin hem de babasının görüşlerine başvurulabilir. Bu durumda gözlem birimi öğrencinin hem annesi hem de babası olur.

17. *Bir araştırmanın evrenini tanımlarken göz önüne alınması gereken ilkeler nelerdir?*
Cevap: Araştırmanın evreni tanımlanırken Göz önüne alınması gereken ilkeler şöyle sıralanabilir:
- Açıklık,
- Kesinlik,
- Amaca uygunluk,
- Örnekleme uygulaması.

18. *Örneklemin aşamaları içerisinde bahsedilen çerçeve kavramı nedir?*
Cevap: Çerçeve sonlu bir evrenin bütün birimlerinin kayıtlı olduğu bir listedir, tablodur veya cetveldir.

Nüfus kayıtları, seçmen kütükleri, tapu ve sicil kayıtları, ticaret ve sanayi odaları üye listeleri, ekonomik büyüklüklerine göre sanayi kuruluşlarının listesi, telefon rehberi, öğrenci kayıt listeleri, su, elektrik abonelik listeleri ve benzeri çerçeve olarak kullanılabilecek araçlardır.

19. *Örneklemeye başlamadan önce ilk araştırılması gereken aşama nedir?*
Cevap: Örneklemeye başlamadan önce amaca uygun bir çerçevenin var olup olmadığı, yoksa sağlanıp sağlanamayacağı öncelikle araştırılmalıdır.

Araştırmaya uygun bir çerçevenin var olması durumunda bu çerçevenin güncel olup olmadığının araştırılması da önemli bir konudur. Dikkat edilmelidir ki çerçeve olmadan ne tam sayım ne de örnekleme yapılabilir.

20. *Çerçeveler oluşturulurken yapılan kapsam hatası nedir?*
Cevap: Bir çerçeve yoksa yeni bir çerçevenin hazırlanması problemi ile karşılaşılır. Bazen tanımlanan evrenin bazı birimleri çerçevede yer almadığı gibi tanımlanan evrenin dışında kalması gereken birimler de çerçevede yer alabilir ya da bazı birimler tekrar tekrar çerçevede yer alabilir. Bu özellikteki çerçevelerde kapsam hatası işlenmiş olur.

21. *Nitel değerlendirmede esas olan faktörler nelerdir?*
Cevap: Esas olan faktörler şöyle sıralanabilir:
- Evrenin homojenliği,
- Araştırmada verilecek kararın önemi,
- Araştırmanın yapısı.

22. *Evrenin homojen veya heterojen yapıda olması örneklem hacmini nasıl etkiler?*
Cevap: Evrenin bütün birimleri ilgilenilen değişken itibariyle aynı değere sahipse bir birimin incelenmesi amaca ulaşmak için yeterlidir.

Ancak birimlerin özellikler bakımından farklılığı arttıkça evreni temsil edebilecek bir örneklem oluşturabilmek için örneklem hacminin de giderek büyümesi gerekir. Yani evrenin homojenliği arttıkça örneklem hacminin azalması beklenir.

23. *Araştırmada verilecek kararın önemi örneklem hacmini nasıl etkiler?*
Cevap: Önemli kararlar için olabildiğince çok veriye ve ayrıntılı bilgiye ihtiyaç vardır. Bu gibi durumlar büyük hacimli bir örneklem üzerinde araştırma yapmayı gerekli kılar. Ancak örneklem hacmi arttıkça maliyet ve gereksinim duyulan zaman ve nitelikli personel sayısı da artar.

Burada dikkat edilmesi gereken husus, bir yandan küçük hacimli örneklem oluşturmak suretiyle bu örneklemin evreni temsil etmesi bakımından yetersiz kalmasını engellemek, diğer taraftan da gereksiz yere çok büyük hacimli örneklem seçerek zaman ve maliyet yönünden kayba uğramamak için uygun büyüklükte bir örneklem hacmini belirlemektir.

Örneklem hacmi arttıkça örnekleme seçilecek her yeni birimin alınacak kararın, yapılacak tahminin doğruluğuna katkısının azalabileceğini dikkate almak gerekir.

24. *Araştırmanın yapısı örneklem hacmini nasıl etkiler?*
Cevap: Uygulamada genellikle nitel araştırmalarda küçük hacimli örneklemlerde, nicel araştırmalarda ise örneğin betimsel araştırmalarda daha büyük hacimli örneklemlerle çalışılır.

Ayrıca araştırmalarda değişken sayısı arttıkça örneklem hacminin arttırılması bilginin niteliği açısından ihtiyaç olur. Örneğin çok değişkenli analiz teknikleri ve yöntemlerinin kullanıldığı araştırmalarda örneklem hacmi büyük olmalıdır.

ÖRNEKLEME YÖNTEMLERİ

25. *Olasılıklı olmayan örnekleme yöntemleri nelerdir?*
Cevap: Olasılıklı olmayan örnekleme yöntemleri şöyle sıralanabilir:

- Kolayda örnekleme,
- Yargısal örnekleme,
- Kota örneklemesi,
- Kartopu örneklemesi.

26. *Olasılıklı örnekleme yöntemleri nelerdir?*
Cevap: Olasılıklı örnekleme yöntemleri şöyle sıralanabilir:

- Basit rassal örnekleme,
- Tabakalı örnekleme,
- Sistematik örnekleme,
- Tek ve çok anlamlı örnekleme.

27. *Olasılıklı olmayan örnekleme kavramını açıklayınız?*
Cevap: Araştırmayı planlayan ya da örnekleme uygulamasını yapan kişi ya da grubun istekleri ve değer yargıları örnekleme seçilecek birimlerin ve örneklem hacminin belirlenmesinde etkili oluyorsa yapılan örnekleme olasılıklı olmayan örneklemedir. Bu tür örnekleme yöntemlerinde, örneklem için birim seçiminde keyfi seçim usulünün uygulandığı örnekleme yöntemleridir. Örneklem oluşturulurken tanımlanan evreni oluşturan birimler arasında fark gözetilir ve bütün birimlere, bilinen bir olasılıkla seçilme şansı verilmezse yapılan seçim keyfi seçimdir.

28. *Olasılıklı olmayan örnekleme yöntemlerinin başarısı nelere bağlıdır?*
Cevap: İlgili örnekleme yöntemlerinin başarısının nelere bağlı olduğu şöyle sıralanabilir:

- Örnekleme uygulamasını yürüten kişi ya da grubun araştırma konusuyla ilgili deneyimine,
- Tanımlanan evrenin özellikleri hakkındaki önsel bilgilere,
- Bu evrenin ilgilenilen özelliklerinin homojenliğine.

29. *Olasılıklı olmayan örnekleme yöntemlerinin ortak özellikleri nelerdir?*
Cevap: İlgili örnekleme yöntemlerinin ortak özellikleri şöyle sıralanabilir:

- Örneklem için birim seçimi keyfidir.
- Örneklem hacmi keyfi olarak belirlenir.
- Örneklemden hesaplanan istatistikler evren parametreleri hakkında genelleme amacıyla kullanılmaz.

30. *Kolayda örnekleme kavramını açıklayınız?*
Cevap: Araştırma konusu ile ilgili ve kolayca ulaşılabilir olan birimlerden oluşturulan örnekleme yöntemidir. Araştırma konusu ile ilgili olan ve doğru yerde, doğru zamanda bulunan birimler arasından keyfi olarak birimler seçiliyorsa yapılan örneklemeye ***kolayda örnekleme*** denir. Kolayda örnekleme gönüllülük esasına göre katılan birimlerden oluşur.

31. *Yargısal örnekleme kavramını açıklayınız?*
Cevap: Örneklemin araştırmacının ya da örneklemeyi yapan kişinin kişisel arzu, düşünce ve deneyimine göre seçilmiş olduğu örneklemedir.

32. *Yargısal örneklemenin kolayda örnekleme yönteminden farkı nedir?*
Cevap: Yargısal örneklemenin kolayda örneklemeden farkı örnekleme birimi seçimi için araştırmacının uzman fikirleriyle belirlediği ölçütler kullanması ve bu ölçütlerin temsili bir örneklem oluşturacak ölçütler olduğuna inanıyor olmasıdır.

33. *Yargısal örneklemenin kullanıldığı alanlar nelerdir?*
Cevap: Yargısal örneklemenin kullanıldığı alanlar şöyle sıralanabilir:

- Pazarlama araştırması,
- Kamuoyu araştırmaları,
- Biyolojik araştırmalar.

34. *Kota örneklemesi kavramını açıklayınız?*
Cevap: Evrenin belirli özelliklere göre tabakalara ayrılarak, tabakaların içerdiği gözlem hacimlerinin toplam gözlem hacmindeki oranı gözetilerek her tabakadan bu oranda gözlem seçilerek yapılan bir örnekleme yöntemidir.

35. *Kota örneklemesi hangi durumda tercih edilmelidir?*
Cevap: Tanımlanan sonlu evren heterojen özelliklere sahip birimlerden oluşuyorsa kota örnekleme tercih edilmelidir.

36. *Kota örneklemesinin başarıyla uygulanabilmesi hangi koşullara bağlıdır?*
Cevap: Bu yöntemin başarıyla uygulanabilmesi için tanımlanan sonlu evrenle ilgili bir çerçevenin var olması, ilgili evrenin homojen veya heterojen özelliğe sahip olup olmadığının sorgulanabilmesi için evren hakkında öncül bilgilere sahip olunması, evrenin heterojen olduğuna karar verilmiş ise hangi kritere göre heterojen birimlerden oluşan bu evrenin homojen birimlerden oluşacak tabakalara ayırmada kullanılacak kriterin belirlenmesi ve tabaka hacimlerinin bilinmesi gerekir.

37. *Kartopu örneklemesi yöntemini açıklayınız?*
Cevap: Özellikle bir çerçevenin var olmaması durumunda ya da oluşturulmasının imkansız olduğu durumlarda faydalı bir örneklemedir. Bu yöntemde örnekleme süreci tanımlanan evrende yer alan bir bireyin genellikle rassal olarak seçilmesiyle başlar. Belirlenen bu birey örnekleme giren birinci birimdir. Bu bireyden aynı evren tanımında yer alan tanıdığı bir bireyin olup olmadığı öğrenilir. Varsa bu bireye ulaşılır. Böylece örneklemde yer alacak ikinci birime ulaşılmış olur. Benzer şekilde bu süreç, referanslarla keyfi olarak belirlenen hacimde örnekleme ulaşıncaya kadar sürdürülür.

ÖRNEKLEME DAĞILIMLARI

38. *Örnekleme dağılımı kavramını açıklayınız?*
Cevap: Bir kitleden seçilen rassal örneklemler için hesaplanan istatistiklerin dağılımına ***örnekleme dağılımı*** denir.

39. *Ortalamanın örnekleme dağılımının nasıl hesaplandığını açıklayınız?*
Cevap: Bir kitleden seçilen rassal örneklemler için hesaplanan ortalama değerlerinin dağılımı bulunarak hesaplanan dağılıma ***ortalamanın örnekleme dağılımı*** denir.

40. *Örneklem olarak seçilen gözlem değerlerinin sayısının artması örneklemenin ortalama dağılımını nasıl etkiler?*
Cevap: Örneklem hacmi arttıkça ortalamanın örnekleme dağılımının ortalaması evren ortalamasına yaklaşır. Örneklem hacmi yeterli büyüklüğe ulaştığında ortalamanın örnekleme dağılımı normal olur.

41. *Ortalamanın dağılımının standart hatası nedir?*
Cevap: Kitleden çekilen örneklemler üzerinden hesaplanan ortalama değerlerinin standart hatasıdır. Ortalama dağılımının standart hatası ortalamanın örnekleme dağılımının değişkenliğini gösterir. Yani mümkün örneklem ortalamalarının kitle ortalamasından farklarının ortalama ölçüsüdür. Ortalama dağılımının standart hatasının karesi ortalama dağılımının varyansını ifade eder.

42. *Merkezi limit teoremini açıklayınız?*
Cevap: Örneklem hacmi büyüdükçe ortalamanın örnekleme dağılımının normal dağılıma yaklaşmasıdır. Örneklem hacmi için yeterli büyüklü kesin olmamakla birlikte uygulamada en az 30 birim olarak kabul edilir.

43. *Örneklem oranı kavramı nedir, bir örnek ile açıklayınız?*
Cevap: Örnekleme planlarında ele alınan evrenin araştırılmak istenen özelliklerinin bazıları iki sonuçlu olmaktadır. Örneğin bir fabrikada üretilen ürünler, hatalı ya da hatasız ürün, bir fakültedeki öğrenciler, başarılı ya da başarısız öğrenci olmak üzere iki grupta toplanabilir. Bu iki sonuçtan birinde örneğin A sonucunda yer alan birimlerin oranıyla ilgilenilebilir. Bu durumda evren oranı evrenin birimleri içindeki ilgilenilen türden özelliğe sahip olanların oranı biçiminde tanımlanır.

44.*Oranların örnekleme dağılımı kavramını açıklayınız?*
Cevap: İki sonuçlu bir evrenden, mümkün bütün n hacimli basit rassal örneklemlerin seçildiğini ver her örneklem için p oranının hesaplandığı varsayıldığında p oranlarından oluşan bir dağılım elde edilir. Örnekleme planlarında, tanımlanan evrenden rassal olarak n hacimli sadece tek bir örneklem oluşturulur ve bu örneklem için p oranı hesaplanır. P rassal değişkeninin çekilmesi mümkün bütün n hacimli örneklemlerde aldığı değerlerin dağılımına ***oranların örnekleme dağılımı*** denir.

45. *Oranların örnekleme dağılımının gösterdiği dağılım şekilleri nasıl belirlenir?*
Cevap: Oranların örnekleme dağılımının gösterdiği dağılım şekilleri aşağıdaki şekilde belirlenebilir:

- $E(p)<0.5$ ise sağa çarpık,
- $E(p)=0.5$ ise simetrik,
- $E(p)>0.5$ ise sola çarpık bir dağılım gösterir.

E(p) değeri 0 veya 1'e yaklaşırken dağılımın çarpıklığı artar.

46. *Merkezi limit teoremine göre oranların örnekleme dağılımı hangi koşulda normal dağılıma yaklaşır?*
Cevap: Merkezi limit teoremine göre rassal örneklem hacmi en az 30 birim olması ve evren oranının 0 ya da 1'e yakın değerler almaması koşuluyla oranların örnekleme dağılımı normal dağılıma yaklaşır. Bu koşulları sağlayan oranların örnekleme dağılımı ile ilgili problemlerin çözümlerinde normal dağılımın özelliklerinde yararlanılır.

47. *Örneklem hacminin belirlenmesinde kullanılan nicel yöntemler nelerdir?*
Cevap: İlgili nicel yöntemler şöyle sıralanabilir:

- Karşılanabilecek maliyeti esas alan yöntem,
- Kabul edilebilir hata düzeyini esas alan yöntem.

İST203U-İSTATİSTİK

Ünite 7: İstatistiksel Tahminleme ve Karar Alma

GİRİŞ

1. *Tam sayım yapılamayıp, örneklemeye başvurulunca karşılaşılan problemler nelerdir?*

Cevap: Tam sayım yapılamayıp, örneklemeye başvurulunca araştırmacılar iki problem ile karşılaşırlar:

- İstatistiksel tahminleme,
- İstatistiksel karar verme.

İSTATİSTİKSEL TAHMİNLEME

2. *Tahminleme nedir?*

Cevap: Tahminleme, tanımlanan evrenden seçilen rassal örneklemden hesaplanan istatistikler yardımıyla bu evrenin uyduğu dağılımın parametre değerlerini araştırmaktır.

3. *Bir tahminleme sürecinde izlenmesi gereken aşamalar nelerdir?*

Cevap: Bir tahminleme sürecinde aşağıdaki aşamalar izlenir:

- Tanımlanan evrenden önceden belirlenen *n* hacimli rassal bir örneklem seçilir.
- Örneklemdeki birimler üzerinden veriler (gözlem değerleri) derlenir.
- Derlenen bu gözlem değerleri kullanılarak tahminlenecek parametre için bilgi üretecek istatistikler hesaplanır.
- Parametre için bilgi üreten istatistiğin örnekleme dağılımının özelliklerinden yararlanılarak parametre değeri tahminlenir.

4. *"Tahminleyici" ile "tahmin" kavramları açıklayınız ve aralarındaki farkı belirtiniz.*

Cevap: Hem örneklem için hem de evren için bilgi üreten istatistiğe ilişkin formülasyona "tahminleyici", örneklem gözlem değerinin bir tahminleyiciye uygulanmasıyla hesaplanan değere ise ***"tahmin"*** adı verilir. Tahminleyici, tahminin nasıl yapılacağını gösterir. Tahmin ise sayısal bir değerdir.

İSTATİSTİK TAHMİNLEME TÜRLERİ

5. *Nokta tahminlemesi nedir?*

Cevap: Bir rassal örneklemden hesaplanan istatistiğin değerini ilgili evren parametresi $\hat{\theta}$ değerine eşit kabul eden tahminleme sürecine ***nokta tahminlemesi*** denir.

6. *Nokta tahminlemesi sürecinin aşamaları nelerdir?*

Cevap: Nokta tahminlemesi sürecinin aşamaları şöyle sıralanabilir:

- Tanımlanan evrenden, belirlenen n birimlik (hacimli) bir basit rassal örneklem seçilir.
- Bu örneklemdeki birimler üzerinden ilgilenilen değişken itibariyle veriler derlenir.
- Bu veriler kullanılarak tahminlenecek parametre θ için tahminleyici olan istatistik hesaplanır.
- Son olarak bu istatistiğin değerinden yararlanarak θ parametresi için tahminleme yapılır.

7. *Evren aritmetik ortalaması μ'nün nokta tahmini nedir?*

Cevap: Bir rassal değişken olan $\bar{X}$ 'nin değerinin bu istatistiğin bilgi ürettiği μ'nün değerine eşit olan tahminlemeye μ'nün nokta tahminlemesi, hesaplanan $\bar{X}$ değerine ***μ'nün nokta tahmini*** denir.

8. *Bir araştırmacı ilaç dağıtım şirketi çalışanlarının yıllık ortalama seyahat etme sürelerini tahmin etmek istemektedir. Bu amaçla rassal olarak 100 ilaç dağıtıcısı seçilmiş ve yıllık ortalama seyahat etme sürelerinin 386 saat ve standart sapmanın 72 saat olduğu belirlenmiştir. Yıllık seyahat etme süresinin nokta tahminlemesini elde ediniz.*

Cevap: 100 birimlik rassal örneklemin aritmetik ortalaması, evren ortalamasının bir nokta tahmini olduğundan, aylık ortalama seyahat etme süresine ilişkin tahmin E ($\bar{X}$) = μ = 386 saat olacaktır.

9. *Evren oranı Π'nin nokta tahminlemesi nedir?*

Cevap: Evren oranı Π'ye ilişkin nokta tahminlemesi, bir rassal örnekleme planında oluşturulan *n* hacimli örneklem için, r bir binom rassal değişkeni olmak üzere, hesaplanan $p = \frac{r}{n}$ oranının değeri Π'ye eşit olan bir tahminleme sürecidir.

10. *Bir elektrikli ısıtıcı üreticisi yeni ürettikleri modelden memnun olan müşterilerinin oranını tahmin etmek istiyor. Bu amaçla yeni elektrikli ısıtıcıyı kullananlar arasından 80 kişi seçiliyor ve bunların 56'sının yeni ısıtıcı modelinden memnun oldukları belirleniyor. İstenen tahmini nokta tahmini olarak yapınız.*

Cevap: *n* = 80 kişi ve *r* = 56 kişi olduğundan,

$$p = \frac{r}{n} = \frac{56}{80} = 0.70 \text{ olur.}$$

Örneklem oranı E(p) = Π = 0.70 değeri evren oranı Π'nin nokta tahminidir.

11. *Nokta tahminlemesi neden sınırlı bir tahminlemedir?*

Cevap: Güvenilir tahmin, tanımlanan evrenden seçilen aynı hacimli farklı örneklemlerde büyük ölçüde farklılık göstermeyen tahmindir. Nokta tahminlemesi *tahminin güvenilirliği hakkında bilgi veremediği için* sınırlı bir tahminlemedir.

12. *Aralık tahminlemesi nedir?*

Cevap: Bir parametrenin örneklem istatistiğine göre, örneklemenin planlama aşamasında belirlenen bir olasılığa (güven düzeyine) göre simetrik bir aralıkta belirlenmesi çalışmasına ***aralık tahminlemesi*** denir.

13. *Aralık tahminlemesinin genel gösterimi nasıldır?*

Cevap: Aralık tahminlemesinin genel gösterimi;

- A : Alt sınır ve
- Ü : Üst sınır olmak üzere aşağıdaki şekildedir:

$$A \leq \theta \leq Ü$$

14. *Güven düzeyi nedir?*
Cevap: Aralık tahminlemesi sürecinde, araştırmacı tarafından önceden belirlenen ***güven düzeyi (G.D.)*** doğru aralık tahminlemesinin yapıldığı parametre değerini kapsayan güven aralığının tahminlendiği olasılığı ifade eder.

15. *Güven düzeyi nasıl ifade edilir?*
Cevap: Güven düzeyi, α : Anlamlılık düzeyi olmak üzere aşağıdaki şekilde ifade edilir:

G.D.= 1-α

16. *Güven düzeyi değeri büyüdükçe nasıl bir sonuçla karşılaşılır?*
Cevap: G.D.= 1-α değeri büyük seçilirse tahminlenen aralığın θ'yı kapsayan bir aralık olma olasılığı artmış, fakat tahminlerin güvenilirliği azalmış olur.

17. *Bir aralık tahminlemesi sürecinde, güven sınırlarının tahminlenebilmesi için hangi adımlar izlenir?*
Cevap: Güven sınırlarının tahminlenebilmesi için şu adımlar izlenir:
- G.D.=1-α belirlenir.
- Hacmi n olan bir basit rassal örneklem seçilir. Tahminlenecek parametre için, bilgi üretecek $\hat{\theta}$ istatistiği hesaplanır.
- Bu istatistiğin dağılımıyla ilgili bilgilerden yararlanarak güven aralığı oluşturulur.

18. *Uygulamalarda genellikle tercih edilen güven düzeyi değerleri hangileridir?*
Cevap: Genellikle tercih edilen güven düzeyi %95 ve %99'dur.

19. *Bir hipotez testinde güven düzeyi %99 olarak belirlenirse anlamlılık düzeyi ne olur?*
Cevap: G.D.=1-α olduğundan α=1-G.D. olmalıdır. Dolayısıyla anlam düzeyi α=1-0.99=0.01 olur.

19. *Bir hipotez testinde güven düzeyi %99 olarak belirlenirse anlamlılık düzeyi ne olur?*
Cevap: G.D.=1-α olduğundan α=1-G.D. olmalıdır. Dolayısıyla anlam düzeyi α=1-0.99=0.01 olur.

20. *Bir hipotez testinde anlam düzeyi 0.05 ise güven düzeyi hangi değer olarak belirlenmiştir?*
Cevap: G.D.=1-α olduğundan G.D.=1-0.05=0.95'dir. Dolayısıyla güven düzeyi %95 olarak belirlenmiştir.

21. *Hesaplanan $\hat{\theta}$ istatistiğinin dağılımı normal ise θ parametresi için aralık tahminlemesinin sınırları hangi eşitlikler ile ifade edilir?*
Cevap: İlgili aralık tahminlemesinin sınırları aşağıdaki şekilde ifade edilir:
- $\hat{\theta} - z_{\frac{\alpha}{2}}.\sigma_{\hat{\theta}} < \theta < \hat{\theta} - z_{\frac{\alpha}{2}}.\sigma_{\hat{\theta}}$ (σ bilindiğinde)
- $\hat{\theta} - z_{\frac{\alpha}{2}}.s_{\hat{\theta}} < \theta < \hat{\theta} - z_{\frac{\alpha}{2}}.s_{\hat{\theta}}$ (σ bilinmediğinde)

Eşitliklerde;
- $\hat{\theta}$: Örneklem istatistiğinin değerini,
- $z_{\frac{\alpha}{2}}$: belirlenen güven düzeyi için standart dağılım tablo değerini,
- $\sigma_{\hat{\theta}}$: standart hatayı ve
- $s_{\hat{\theta}}$: σ bilinmediğinde (σ yerine onun yansız tahmini s kullanıldığında) standart hata tahminini göstermektedir.

22. *Aralık tahminlemesinde güven aralığının mümkün olduğu ölçüde dar tutulması arzu edilir. Neden?*
Cevap: Çünkü dar aralığın sınırları parametre değerine daha yakındır.

23. *Aralık tahminlemesinde güven aralığının genişliği nelere bağlıdır?*
Cevap: Güven düzeyine ve örneklem hacmine bağlıdır.

24. *Evren aritmetik ortalaması μ için 1-α güven sınırlarının ya da güven aralığının belirlenmesi işlemlerine ne denir?*

Cevap: Soruda verilen ilgili işlemlere ***μ'nün aralık tahminlemesi*** denir.

25. *Evren aritmetik ortalaması μ'nün aralık tahminlemesi sürecinde hangi aşamalar izlenir?*

Cevap: İlgili tahminleme sürecinde şu aşamalar izlenir:
- $G.D. = 1 - \alpha$ belirlenir.
- Hacmi n olan bir rassal örneklem seçilir.
- Bu örneklem için $\bar{X}$ ve s istatistikleri hesaplanır.
- $\bar{X}$'nın standart hatası $\sigma_{\bar{x}}$ hesaplanır ya da $s_{\bar{x}}$ tahminleyicisi yardımıyla tahmin edilir.
- $\bar{X}$'nın dağılım şekliyle ilgili bilgilerden yararlanarak $\bar{X}_A < \mu < \bar{X}_Ü$ güven aralığı tahminlenir.

26. *Büyük örneklem hacmi için yeterli büyüklük ne olmalıdır?*
Cevap: Büyük örneklem hacmi için yeterli büyüklük $n \geq 30$ birim olmalıdır.

27. *Rassal örneklem hacmi yeterli büyüklükte ise $\bar{X}$'nın örnekleme dağılımı hakkında ne söylenebilir?*
Cevap: Örneklem hacmi yeterli büyüklükte ise, evren dağılım şekli ne olursa olsun, $\bar{X}$'nın örnekleme dağılımı, ortalaması μ ve varyansı $\frac{\sigma^2}{n}$ olan normal dağılıma uyar.

27. *z standart rassal değişkeni nasıl hesaplanır?*
Cevap: z standart rassal değişkeni aşağıdaki formül ile hesaplanır:

$$z = \frac{\bar{X} - \mu}{\sigma_{\bar{X}}}$$

28. *Bilinmeyen evren parametresi μ için güven aralığı nasıl oluşturulur?*
Cevap: İlgili güven aralığı aşağıda verilmiştir:

$$\bar{X} - z_{\frac{\alpha}{2}}\sigma_{\bar{x}} < \mu < \bar{X} + z_{\frac{\alpha}{2}}\sigma_{\bar{x}}$$

(*Dikkat:* Burada z değeri, araştırmacı tarafından başlangıçta belirlenen $1-\alpha$ güven düzeyine bağlı olarak standart normal eğri alanları tablosundan bulunur.)

29. *%99 güven düzeyi için $z_{\frac{\alpha}{2}}$ değeri nedir?*

Cevap: α=1-0.99=0.01'dir.

$$z_{\frac{\alpha}{2}} = z_{\frac{0.01}{2}} = z_{0.005} = 2.576$$

30. *%95 güven düzeyi için $z_{\frac{\alpha}{2}}$ değeri nedir?*

Cevap: α=1-0.95=0.05'dir.

$$z_{\frac{\alpha}{2}} = z_{\frac{0.05}{2}} = z_{0.025} = 1.960$$

31. *Örneklemin seçiminde iadesiz seçim uygulanıyor ya da örnekleme oranı $n/N \geq 0,05$ olduğunda ortalamanın standart hatası ($\sigma_{\bar{x}}$) nasıl hesaplanır?*

Cevap: $\bar{X}$'nın standart hatası, örneklemin seçiminde iadesiz seçim uygulanıyor ya da örnekleme oranı $n/N \geq 0,05$ olduğunda aşağıdaki formül ile hesaplanır:

$$\sigma_{\bar{x}} = \frac{\sigma}{\sqrt{n}}.\sqrt{\frac{N-n}{N-1}}$$

32. *Ortalamanın standart hatasının tahmini nasıl hesaplanır?*
Cevap: Ortalamanın standart hatasının tahmini ($s_{\bar{x}}$); örnekleme birim seçiminde iadeli seçim uygulanırsa ya da örnekleme oranı $n/N < 0,05$ olursa ya da tanımlanan evren sonsuz olursa aşağıdaki formül ile hesaplanır:

$$s_{\bar{x}} = \frac{s}{\sqrt{n}}$$

Buna karşılık birim seçimi iadesiz yapılır ya da örnekleme oranı $n/N \geq 0,05$ olursa aşağıdaki formül ile hesaplanır:

$$s_{\bar{x}} = \frac{\sigma}{\sqrt{n}}.\sqrt{\frac{N-n}{N-1}}$$

33. *Rassal örneklem hacmi yeterli büyüklükte olduğunda 1-α güven düzeyi için, μ'nün güven aralığı nasıl bulunur?*
Cevap: İlgili güven aralığı aşağıda verilmiştir:

$$\bar{X} - z_{\frac{\alpha}{2}}s_{\bar{x}} < \mu < \bar{X} + z_{\frac{\alpha}{2}}s_{\bar{x}}$$

34. *Konserve bezelye üreten bir fabrikadan, üretim sırasında 64 konserve kutusundan oluşan rassal bir örneklem seçilmiştir. Bu konserve kutularının ortalama ağırlığı 492 gr. ve standart sapma 12 gr. olarak belirlenmiştir. Üretilen konservelerin ortalama ağırlığını %95 güven düzeyinde tahmin ediniz.*
Cevap: $n = 64$ konserve

($n > 30$'dur. Ayrıca evren hacmi sonsuzdur.)

$\bar{X} = 492$ gr., $s = 12$ gr.

G.D.=1-α=0.95 dolayısıyla α=0.05'dir.

$$z_{\frac{\alpha}{2}} = z_{\frac{0.05}{2}} = z_{0.025} = 1.96$$

$$s_{\bar{x}} = \frac{s}{\sqrt{n}} = \frac{12}{\sqrt{64}} = 1.5 \text{ gr.}$$

$$\bar{X} - z_{\frac{\alpha}{2}}s_{\bar{x}} < \mu < \bar{X} + z_{\frac{\alpha}{2}}s_{\bar{x}}$$

$$492-(1.96)(1.5) < \mu < 492-(1.96)(1.5)$$
$$489.06 < \mu < 494.94$$

Üretilen konserve kutularının ortalama ağırlığı %95 güven ile 489.06 gr. ile 494.94 gr. arasındadır.

35. *Örneklem hacmi küçük ($n < 30$) olduğu zaman nasıl bir durumla karşılaşılır?*
Cevap: Örneklem hacmi küçük ($n < 30$) olduğu zaman, örneklem ortalamalarının standart değerleri gibi normal dağılıma sahip olmaz.

Bu durumda μ için aralık tahminlemesi, tanımlanan evrenin normal dağılıma sahip olup olmadığının bilinmesine bağlıdır.

36. *Normal dağılan bir evrenden, rassal olarak seçilebilecek birbirinden farklı $n < 30$ birim hacimli mümkün bütün örneklemlerin seçildiği ve her örneklem için $\bar{X}$'ların ve $\frac{(\bar{X}-\mu)}{s_{\bar{x}}}$ standart değerlerinin hesaplandığını durumlarda değer aralığı $-\infty$ ile $+\infty$ olan $\frac{(\bar{X}-\mu)}{s_{\bar{x}}}$ istatistiğinin dağılımı ne olur?*

Cevap: *s.d = n - 1* serbestlik derecesinde *t* dağılımı adı verilen sürekli bir dağılıma uyar.

37. *t dağılımı nasıl bir dağılımdır?*
Cevap: *t* dağılımı ortalaması sıfır olan tek modlu ve simetrik bir dağılımdır. Bu dağılımın şekli normal dağılımın şekline benzer fakat değişkenliği daha fazladır.

Örneklem hacmi artarken serbestlik derecesi *s.d= n-1* büyür, $s_{\bar{x}}$'nin kullanılması nedeniyle ortaya çıkan değişkenlik küçülür ve *t* dağılımı standart normal dağılıma yaklaşır.

38. *Örneklem hacmi küçük olduğu durumlarda, tanımlanan evren normal dağılıyorsa, μ için güven aralığı nasıl belirlenir?*

Cevap: İlgili güven aralığı aşağıda verilmiştir:

$$\bar{X} - t.s_{\bar{x}} < \mu < \bar{X} + t.s_{\bar{x}}$$

(*Dikkat:* Burada; *t*: $1-\alpha$ güven düzeyinde ve *n-1* serbestlik derecesine bağlı olarak *t* değerleri tablosundan bulunur.)

(*Dikkat:* $s_{\bar{x}}$: Ortalamanın standart hata tahminidir ve $s_{\bar{x}} = \frac{s}{\sqrt{n-1}}$ tahminleyiciyle tahminlenir.)

39. *Evren normal dağılıyorsa, büyük örneklemler söz kunusu iken μ için güven aralığı belirmek ile küçük örneklemler söz kunusu iken μ için güven aralığı belirmek arasındaki fark nedir?*

Cevap: Büyük örneklemler söz konusu ise *z* istatistiği, küçük örneklemler söz konusu ise *t* istatistiği kullanılır.

40. *Bir cep telefonu üreticisi, üretilen cep telefonlarındaki bataryaların bir kez şarj edildikten sonra ortalama ne kadar süre kullanıldığını tahmin etmek istemektedir. Bu amaçla rassal olarak 17 telefon bataryası seçilmiş ve bunların ortalama 135 saat kullanılabildiği ve standart sapmanın 28 saat olduğu belirlenmiştir. %99 güven düzeyi için istenilen tahminlemeyi yapınız.*

Cevap: $n = 17$ batarya ($n < 30$ 'dur.)

$\bar{X} = 135$ saat, $s = 28$ saat

G.D.=1-α=0.99 dolayısıyla α=0.01'dir.

α=0.05 ve s.d.=17-1=16 serbestlik derecesi için

t tablo değeri: $t_{\frac{\alpha}{2};s.d.} = z_{\frac{0.01}{2};16} = z_{0.005;16} = 2.92$

$$s_{\bar{x}} = \frac{s}{\sqrt{n-1}} = \frac{28}{\sqrt{16}} = 7 \text{ saat}$$

$$\bar{X} - t.s_{\bar{x}} < \mu < \bar{X} + t.s_{\bar{x}}$$

$$135-(2.92)(7) < \mu < 492-(2.92)(7)$$
$$114.55 < \mu < 155.48$$

Üretilen cep telefonlarındaki bataryaların bir kez şarj edildikten sonra ortalama kullanım süresi % 99 güvenle 114.553 saat ile 155.447 saat arasında bir değerdir.

41. *Evren dağılımı normal değilse μ'nün güven aralığının tahminlemesi nasıl yapılır?*

Cevap: Evren dağılımı normal değilse verilerin bir matematiksel dönüşümle normale yaklaştırabileceği düşüncesinden hareketle μ'nün güven aralığının tahminlenmesinde, dönüştürülmüş verilere uygulanır. Örneğin logaritmik dönüşüm bu amaçla sıkça kullanılır. Çünkü dönüştürülmüş logaritmik veriler, dönüştürülmemiş verilere göre daha az çarpıktır.

42. *Büyük örneklemler ($n \geq 30$) söz konusu is, evren oranı Π'nin aralık tahminlemesi nasıl yapılır?*

Cevap: Büyük örneklemler söz konusu iken örneklem oranı (*p*)'nin örnekleme dağılımı, yaklaşık normal dağılım gösterir. Aynı zamanda bu dağılımın ortalaması p, standart hatası;

$$s_p = \sqrt{\frac{p(1-p)}{n}} \text{'dir.}$$

Rassal örneklem hacmi yeterli büyüklükteyse, örneklem oranlarının standart değerlerinin ($z = \frac{p-\Pi}{s_p}$), standart normal dağılıma sahip olduğu söylenebilir ve evren ortalaması aralık tahmininde olduğu gibi $1-\alpha$ güven düzeyinde evren oranı Π 'nin aralık tahmini şu şekilde hesaplanır:

$$p - z_{\frac{\alpha}{2}} s_p < \Pi < p + z_{\frac{\alpha}{2}} s_p$$

43. *Bir bölgede yaşayan kişiler arasından A gazetesi okuru olanların oranı belirlenmek isteniyor. Bu amaçla rassal olarak 200 kişiden oluşan bir örneklem seçiliyor ve seçilen 200 kişi arasından 58 kişinin A gazetesi okuru olduğu belirleniyor. Bu bölgedeki A gazetesi okurlarının oranını %95 güven düzeyinde tahmin ediniz.*

Cevap: $n = 200$ kişi ($n > 30$ 'dur.)

$r = 58$ kişi saat, $s = 28$ saat,

$$p = \frac{r}{n} = \frac{58}{200} = 0.29$$

G.D.=1-α=0.95 dolayısıyla α=0.05'dir.

$$z_{\frac{\alpha}{2}} = z_{\frac{0.05}{2}} = z_{0.025} = 1.96$$

$$s_p = \sqrt{\frac{(0.29)(0.71)}{200}} = 0.032$$

$$p - z_{\frac{\alpha}{2}} s_p < \Pi < p + z_{\frac{\alpha}{2}} s_p$$

$$0.29-(1.96)(0.032) < \mu < 0.29-(1.96)(0.032)$$
$$0.22728 < \mu < 0.35272$$

Bu bölgedeki A gazetesi okurlarının oranı %95 güvenle, %22.73 ile %35.27 arasında bir değer olarak belirlenir.

İSTATİSTİKSEL HİPOTEZ VE İSTATİSTİKSEL HİPOTEZ TESTİ

44. *İstatistiksel hipotez nedir?*
Cevap: İstatistiksel hipotez, bir araştırmada ilgilenilen bir ya da daha fazla parametrenin değeri hakkında ileri sürülen ve doğruluğu, geçerliliği bu parametre hakkında bilgi üreten istatistikten ve bu istatistiğin örnekleme dağılımıyla ilgili bilgilerden yararlanarak araştırılabilen önermelerdir.

45. *İsatistiksel hipotez testinin konusu nedir?*
Cevap: İstatistiksel hipotezler bir ya da daha fazla evren parametre değeriyle ilgili olabilirler.

46. *İstatistiksel hipotezleri diğer hipotezlerden ayıran özellik nedir?*
Cevap: İstatistiksel hipotezleri diğer hipotezlerden ayıran özellik, bu hipotezlerin bir frekans dağılımına ait olmasıdır.

HİPOTEZ TESTİ TÜRLERİ

47. *Hipotez testleri, ilgilenilen değişkenin ölçümlenmesinde benimsenen ölçeğe bağlı olarak nasıl sınıflandırılırlar?*
Cevap: Hipotez testleri, ilgilenilen değişkenin ölçülmesinde benimsenen ölçeğe bağlı olarak;

- Parametrik hipotez testleri ve
- Parametrik olmayan hipotez testleri şeklinde sınıflandırılırlar.

48. *Parametrik hipotez testlerinde değişkenlerinin ölçülmesinde neden eşit aralıklı ya da oranlı ölçek kullanılmaktadır?*
Cevap: Çünkü bu iki ölçekte de elde edilen veriler üzerinden aritmetik işlemler yapmak mümkündür.

49. *En önemli parametrik testler hangileridir?*
Cevap: En önemli parametrik testler;

- *"z"* testi ve
- *"t"* testidir.

HİPOTEZ TESTİ SÜRECİNİN ADIMLARI

50. *Hipotez testi sürecinin adımları nelerdir?*
Cevap: Hipotez testi sürecinin adımları şöyle sıralanabilir:

- Hipotezlerin ifade edilmesi,
- Anlamlılık düzeyinin belirlenmesi,
- Verilerin derlenmesi,
- Test istatistiğinin seçilmesi,
- İstatistiksel kararın verilmesi,
- Probleme ilişkin kararın verilmesi.

51. *Sıfır hipotezi (H_0) nedir?*
Cevap: Sıfır hipotezi, test edilecek hipotezdir. H_0 hipotezinde test süreci tamamlanıncaya kadar $\theta = \theta_0$ olarak kabul edilir ve $\hat{\theta}$ ile θ parametrelerinin değerleri arasındaki farkın örnekleme hatasından kaynaklanabileceği, bu iki değer arasında gerçekte anlamlı bir farklılık olmadığı ifade edilir.

52. *Sıfır hipotezi (H_0) nasıl ifade edilir?*
Cevap: Sıfır hipotezi aşağıdaki şekilde yazılır:

$$H_0:\ \theta = \theta_0$$

53. *Karşıt hipotez (H_1) nedir?*
Cevap: Karşıt hipotez, H_0 hipotezinin belirli bir olasılıkla reddedilmesi durumunda kabul edilen ve genellikle araştırma hipotezinin ifade edildiği hipotezdir.

Karşıt hipotez, sıfır hipotezini çürüten bir hipotezdir.

54. *Karşıt hipotez (H_1) nasıl ifade edilir?*
Cevap: Hipotez araştırmanın amacına bağlı olarak, üç farklı şekilden biriyle ifade edilir:

- $H_1:\ \theta \neq \theta_0$
  (Evren parametresinin belirlenen (ya da bilinen) θ_0 değerinden, her iki yöndeki (hem küçük hem de büyük yöndeki) anlamlı farklılıkların, test sonucu verilecek kararı etkileyeceği anlamına gelir.)
- $H_1:\ \theta > \theta_0$
  (Verilecek kararın, evren parametre değerinde, sadece büyük yöndeki anlamlı sapmadan etkileneceği anlamına gelir.)
- $H_1:\ \theta < \theta_0$
  (Sadece küçük yöndeki anlamlı farklılığın verilecek kararı etkileyeceği anlamına gelir.)

55. *Hipotez testlerinde kabul ya da red edilen hipotez hangisidir?*
Cevap: Hipotez testlerinde kabul ya da red edilen hipotez, H_0 hipotezidir.

56. *Örnekleme dağılımındaki testin yönünü ya da red bölgesinin yerini belirleyen hipotez hangisidir?*
Cevap: Örnekleme dağılımındaki testin yönünü ya da red bölgesinin yerini belirleyen hipotez, H_1 hipotezidir.

57. *Red bölgesi nedir?*
Cevap: Red bölgesi, H_0 hipotezinin reddedilmesine (H_1 hipotezinin kabul edilmesine) neden olan örneklem istatistiği $\hat{\theta}$ (ya da test istatistiği) ile ilgili değerler aralığıdır.

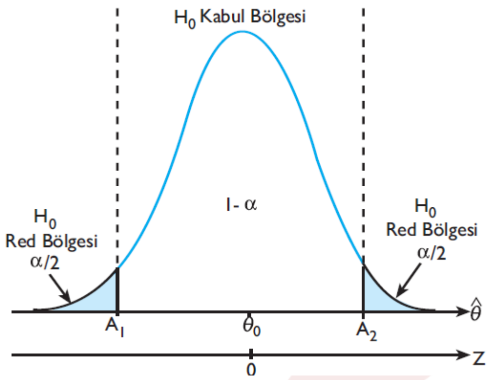
58. *Kabul bölgesi nedir?*
Cevap: Kabul bölgesi, H_0 hipotezinin kabul edilmesine (H_1 hipotezinin reddedilmesine) neden olan örneklem istatistiği $\hat{\theta}$ (test istatistiği) ile ilgili değerler aralığıdır.

59. *Hipotez testleri, H_1 hipotezinin ifade ediliş şekline göre nasıl belirlenir?*
Cevap: Hipotez testleri, H_1 hipotezinin ifade ediliş şekline göre aşağıdaki şekilde belirlenebilir:

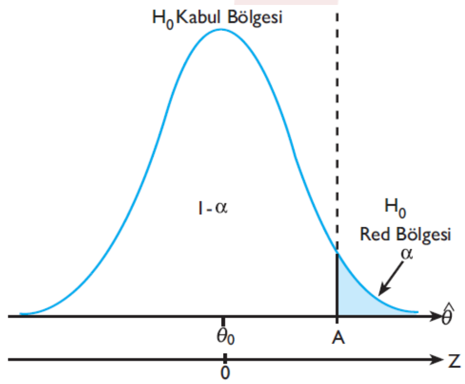
- *İki Yönlü Testlerde Hipotezler:*
  - $H_0:\ \theta = \theta_0$
  - $H_1:\ \theta \neq \theta_0$

- *Tek Yönlü Üst Kuyruk Testlerinde Hipotezler:*
    - H_0: $\theta = \theta_0$
    - H_1: $\theta > \theta_0$
- *Tek Yönlü Alt Kuyruk Testlerinde Hipotezler:*
    - H_0: $\theta = \theta_0$
    - H_1: $\theta < \theta_0$

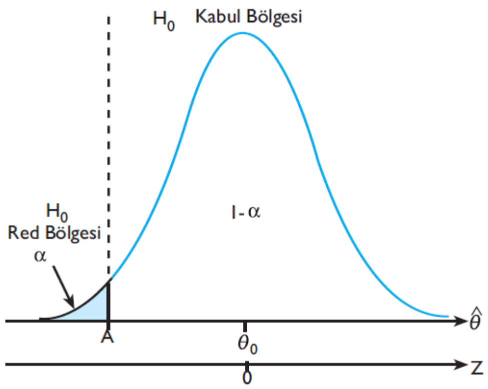
60. *İki yönlü testlerde red ve kabul bölgelerini dağılım grafiği üzerinde gösteriniz.*
Cevap: İlgili grafik aşağıda verilmiştir:



61. *Tek yönlü üst kuyruk testlerinde red ve kabul bölgelerini dağılım grafiği üzerinde gösteriniz.*
Cevap: İlgili grafik aşağıda verilmiştir:



62. *Tek yönlü alt kuyruk testlerinde red ve kabul bölgelerini dağılım grafiği üzerinde gösteriniz.*
Cevap: İlgili grafik aşağıda verilmiştir:



63. *Yorumlama (çıkarsama) hatası nedir?.*
Cevap: Hipotez testlerinde, sıfır hipotezinin yanlışlıkla reddedilmesi ya da kabul edilmesi sonucu işlenen hataya ***"yorumlama (çıkarsama) hatası"*** adı verilir.

64. *Kaç tür yorumlama (çıkarsama) hatası vardır?*
Cevap: İki tür yorumlama hatası vardır:
1. Gerçekte doğru olan sıfır hipotezinin reddedilmesi durumunda işlenen hata,
2. Gerçekte yanlış olan sıfır hipotezinin kabul edilmesi durumunda işlenen hata.

65. *Gerçekte doğru olan sıfır hipotezinin reddedilmesi durumunda işlenen hataya ne ad verilir?*
Cevap: Soruda verilen ilgili hataya ***I. Tip hata (ya da α tipi hata)*** denir.

66. *Testin anlamlılık düzeyi nedir?*
Cevap: α tipi hata işlemenin maksimum olasılığına *"**testin anlamlılık düzeyi**"* denir.

67. *Anlamlılık düzeyinin belirlenmesi işlemi nedir?*
Cevap: Doğru olan sıfır hipotezinin, örneklemden elde edilen bilgilere dayanarak reddedilmesi olasılığını belirleyen α 'nın seçilmesidir. Yapılan bu seçimle birlikte, doğru olan H_0 hipotezinin reddedilme olasılığı, belirlenmiş olur.

68. *Gerçekte yanlış olan sıfır hipotezinin kabul edilmesi durumunda işlenen hataya ne ad verilir?*
Cevap: Soruda verilen ilgili hataya ***II. Tip hata (ya da β tipi hata)*** denir.

69. *α tipi hata ile β tipi hata arasında nasıl bir ilişki vardır?*
Cevap: Örneklem hacmi sabit olduğunda, α tipi hata işlemenin azalması (ya da artması), β tipi hata işleme olasılığının artmasına (ya da azalmasına) neden olur.

70. *Verilerin derlenmesi aşamasında ne yapılır?*
Cevap: Belirlenen evrenden, hangi hacimde bir örneklem seçileceği kararlaştırılır. Daha sonra da ilgili evrenden belirlenen hacimde rassal bir örneklem seçilerek veriler derlenir. Bu veriler kullanılarak, test edilecek parametre hakkında bilgi üreten örneklem istatistikleri hesaplanır.

71. *Test istatistiği nedir?*
Cevap: Test istatistiği, örneklem istatistiği $\hat{\theta}$ ile θ_0 arasındaki farkı standart hata birimiyle ifade eden ölçüdür.

72. *Kritik değer nedir?*
Cevap: Kritik değer, $\hat{\theta}$ istatistiğinin örnekleme dağılımında, red ve kabul bölgelerini birbirinden ayıran bir değerdir.

73. *H_0 hipotezinin reddedilmesi yönündeki kararlar ne anlama gelir?*

Cevap: Örneklem değeri $\hat{\theta}$ ile evren parametresi θ arasında, α anlamlılık düzeyinde anlamlı bir farklılığın var olduğu anlamına gelir.

74. *H_0 hipotezinin kabul edilmesi yönündeki kararlar ne anlama gelir?*

Cevap: Örneklem değeri $\hat{\theta}$ ile evren parametresi θ arasındaki farkın α anlamlılık düzeyinde örnekleme hatasından kaynaklandığı anlamına gelir.

TEK EVREN PARAMETRESİYLE İLGİLİ HİPOTEZ TESTLERİ

75. *Tek evren parametresiyle ilgili yapılan hipotez testlerinde amaç nedir?*

Cevap: Tek evren parametresiyle ilgili hipotezlerin testinde, tanımlanan bir evrenin ilgilenilen bir değişkenine ilişkin önceden belirlenen (ya da bilinen) bir parametre değerinin ($\hat{\theta}$'ın) değişmediği şeklindeki sıfır hipotezi test edilmek amaçlanmaktadır.

Bu testlerde karar verilirken örneklem istatistiğinin değeriyle bu istatistiğin bilgi ürettiği parametrenin bilinen ya da belirlenen değeri θ_0 karşılaştırılır.

76. *Evren ortalamasına ilişkin büyük örneklem testlerinde hangi aşamalar izlenir?*

Cevap: İlgili büyük örneklem testlerinde izlenen aşamalar şöyle sıralanabilir:

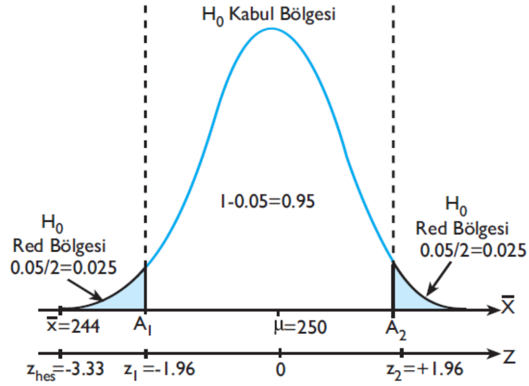
- Örneklem rassal olarak seçilir.
- Örneklem hacminin yeterli büyüklükte ($n \geq 30$) birimden oluştuğu ya da evren normal dağılımlı ve değişkenliğinin biliniyor olması gereklidir.
- H_0: $\mu=\mu_0$ hipotezi, seçilecek bir α anlamlılık düzeyi için test edilir.

77. *Margarin üreten bir fabrikada 250 gr'lık paketler halinde üretim yapılması öngörülmektedir. Margarin paketlerinin ağırlığını kontrol amacıyla rassal olarak 100 paket seçilmiş ve seçilen bu paketler için ortalama ağırlık 244 gr, standart sapma 18 gr olarak saptanmıştır. α=0.05 anlam düzeyinde, paketlerin ağırlığının öngörüldüğü gibi olduğu söylenebilir mi? Karar veriniz.*

Cevap: İlgili adımlar şöyle sıralanabilir:

- ***1. Adım: Hipotezlerin ifade edilmesi:***
  - H_0: $\mu = 250\,gr$
  - H_1: $\mu \neq 250\,gr$
- ***2. Adım: İstatistiksel Test:***
  $n = 100$ paket ($n > 30$)
  Dolayısıyla z testi uygulanmalıdır.
- ***3. Adım: Anlamlılık Düzeyinin Belirlenmesi:***
  α=0.05 olarak belirlenmiştir. Red bölgeleri, ortalamanın örnekleme dağılımının her iki kuyruğunda tanımlandığı için, red bölgelerinin her birinin oransal büyüklüğü,

  α/2 = 0.05/2 = 0.025'dir.
- ***4. Adım: H_0'ın Red Bölgesinin Belirlenmesi:***



- ***5. Adım: Test İstatistiğinin Hesaplanması:***

$$z = \frac{\bar{X}-\mu_0}{s_{\bar{X}}} = \frac{\bar{X}-\mu_0}{\frac{s}{\sqrt{n}}} = \frac{244-250}{\frac{18}{\sqrt{100}}} = -3.33$$

Alt kuyruk bölgesi için kritik değer:

$$z_1 = z_{\frac{\alpha}{2}} = z_{\frac{0.05}{2}} = z_{0.025} = -1.96$$

Üst kuyruk bölgesi için kritik değer:

$$z_1 = z_{\frac{\alpha}{2}} = z_{\frac{0.05}{2}} = z_{0.025} = +1.96$$

$z_{hes} = |3.33| > z_{tab} = 1.96$ olduğundan H_0 hipotezi red edilir, dolayısıyla H_1 kabul edilir.

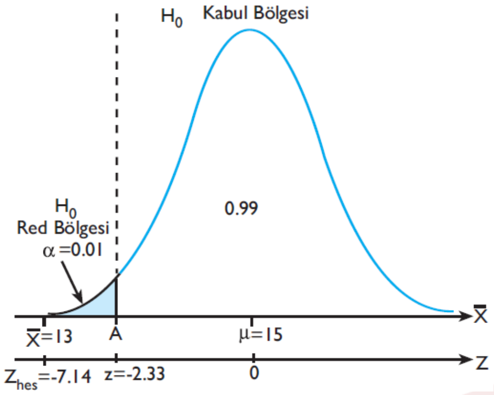
H_0 hipotezinin reddedilmesi, üretilen margarin paketlerinin ortalama ağırlığının 250 gr olmadığını, bu değerden anlamlı bir biçimde farklı olduğunu gösterir.

78. *Bir firmanın geliştirdiği yeni sistemin ortalama paketleme süresini ürün başına 15 dakikanın altına indirdiği iddia edilmektedir. Bu iddiayı araştırmak amacıyla paketleme esnasında rassal olarak seçilen 225 ürünün yeni sistemde ortalama paketlenme süresi 13 dakika ve standart sapması 4.2 dakika olarak belirlenmiştir. Yeni sistemle ilgili iddia hakkında a = 0.01 anlam düzeyinde karar veriniz.*

Cevap: İlgili adımlar şöyle sıralanabilir:

- ***1. Adım: Hipotezlerin ifade edilmesi:***
  Burada verilecek karar, ortalama paketleme süresinin 15 dakikadan az olup olmadığıdır.
  Araştırma hipotezi ise 15 dakikadan az olduğudur. Buna göre hipotezler:
  - H_0: $\mu = 15\,dk$
  - H_1: $\mu < 15\,dk$
- ***2. Adım: İstatistiksel Test:***
  $n = 225$ ($n > 30$)
  Dolayısıyla z testi uygulanmalıdır.

- ***3. Adım: Anlamlılık Düzeyinin Belirlenmesi:*** α=0.01 olarak belirlenmiştir. Red bölgeleri, ortalamanın örnekleme dağılımının alt kuyruğunda tanımlandığı için, red bölgelerinin oransal büyüklüğü $\alpha = 0.01$'dir.
- ***4. Adım: H0'ın Red Bölgesinin Belirlenmesi:***



- ***5. Adım: Test İstatistiğinin Hesaplanması:***

$$z = \frac{\bar{X} - \mu_0}{s_{\bar{X}}} = \frac{\bar{X} - \mu_0}{\frac{s}{\sqrt{n}}} = \frac{13 - 250}{\frac{4.2}{\sqrt{225}}}15 = -7.142$$

Kritik değer:

$$z = z_{\alpha} = z_{0.01} = 2.33$$

$z_{hes} = |-7.142| > z_{tab} = 2.33$ olduğundan H_0 hipotezi red edilir, dolayısıyla H_1 kabul edilir.

H_0 hipotezinin reddedilmesi, bu firmanın geliştirdiği yeni sistemin, ortalama paketleme süresini ürün başına 15 dakikanın altına indirdiğini göstermektedir.

GİRİŞ

1. *Korelasyon analizi nedir?*
Cevap: İki değişken arasındaki ilişkiyi tanımlamaya ve ölçmeye ***korelasyon analizi*** adı verilir.

2. *Regresyon analizi nedir?*
Cevap: Bir veya birden çok değişkenin başka bir değişken üzerindeki ilişkisini açıklamaya ***regresyon analizi*** adı verilir.

3. *Korelasyon ve regresyon analizi arasındaki fark nasıl tanımlanmıştır?*
Cevap: Korelayon analizi ve regresyon analizi tanımları her ne kadar birbirine benzer gözükse de korelasyon analizi ve regresyon analizi bir bozuk paranın iki farklı yüzü gibidirler.

4. *Regresyon terimi ilk olarak ne zaman, kim ve niçin kullanmıştır?*
Cevap: Regresyon terimi 19. yüzyılda İngiliz istatistikçisi Francis Galton tarafından bir biyolojik inceleme için ortaya atılmış ve Galton bu terimi babaların boyları ile oğullarının boyları arasındaki ilişkiyi araştırmak için kullanmıştır.

5. *Regresyon analizi yapabilmek için en az kaç değişkene ihtiyaç duyulmakta ve ihtiyaç duyulan değişken türlerini tanımlayınız?*
Cevap: Regresyon analizinde en az iki farklı değişken tanımlamasına ihtiyaç duyulmaktadır: Bunlar sırasıyla bağımlı ve bağımsız değişken kavramlarıdır. Bağımlı değişken araştırmacının üzerinde çalıştığı değişken olup bu değişken üzerinde meydana gelen değişimlerin ya da bu değişkenin toplam değişkenliğinin açıklanmasına çalışılmaktadır. Bağımsız değişken ya da değişkenler ise ilgilenilen bağımlı değişkende meydana gelen değişim ya da toplam değişkenliğinin üzerinde etkisi olabileceği düşünülen değişken ya da değişkenlerdir.

6. *Tek bağımsız değişken içeren regresyon analizine ne ad verilir?*
Cevap: Araştırmanın tek bağımsız değişken içermesi durumunda yapılan regresyon analizine ***basit doğrusal regresyon analizi*** adı verilmektedir.

7. *Birden fazla bağımsız değişken içeren regresyon analizine ne ad verilir?*
Cevap: Araştırmanın birden fazla bağımsız değişken içermesi durumunda yapılan regresyon analizine ***çoklu doğrusal regresyon analizi*** adı verilmektedir.

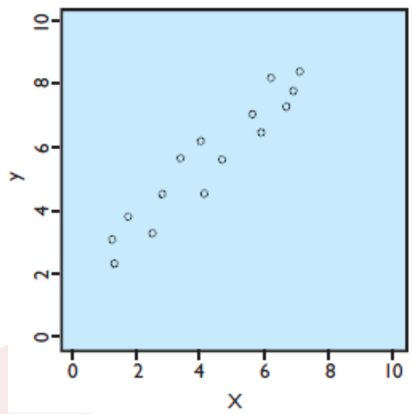
8. *Korelasyon analizinin niçin kullanılır ve regresyon analizinden farklılık nokta nedir?*
Cevap: İlgilenilen iki değişken arasındaki ilişkinin derecesi için korelasyon analizi kullanılır. Korelasyon analizinin regresyon analizinden farklılık gösterdiği nokta, korelasyon analizinin değişkenler arasındaki ilişkinin yalnızca derecesini göstermesidir.

9. *İki değişken arasındaki korelasyonun yüksek olması bu iki değişkeni birini diğerinin nedeni olmasını gösterir mi?*
Cevap: İki değişken arasında yüksek korelasyon olması bu iki değişkenden birinin diğerinin nedeni olabileceğini göstermez.
Korelasyon analizi iki değişken arasındaki nedensellik için kullanılmaz. Nedensellik araştırması için farklı istatistik tekniklerinin kullanılması gerekir.

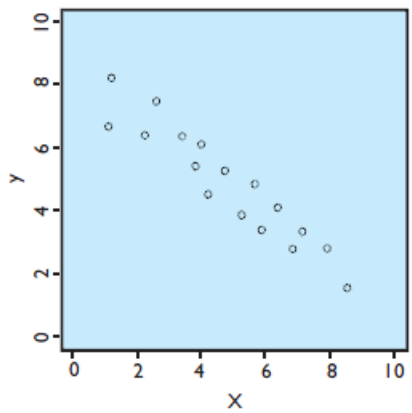
KORELASYON ANALİZİ

10. *Değişkenler arasında var olabilecek ilişkinin derecesini tespit etmek için kullanılan en basit şekli nedir?*
Cevap: Değişkenler arasında var olabilecek ilişkinin derecesinin tespit edilebilmesi amacı ile çeşitli teknikler kullanılabilir. En basit şekli ile iki değişken arasındaki ilişkiyi gözlemlemek için bu değişkenlerin dağılım grafikleri çizilebilir.

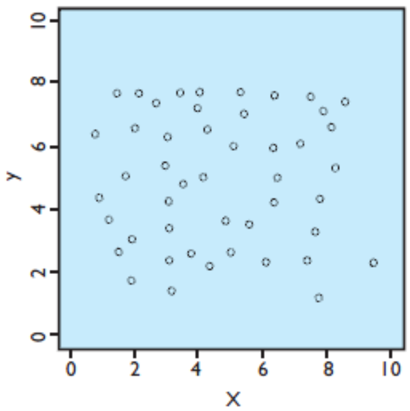
11. *İki değişken arasında pozitif ilişki olduğunda dağılım grafiği nasıl olmalıdır?*
Cevap: Bağımlı y ve bağımsız x değişkenleri arasında pozitif(+) ilişki olduğunda bu değişkenlerin dağılım grafikleri aşağıdaki gibi olur:



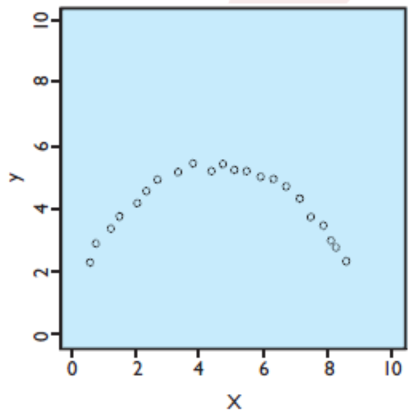
12. *İki değişken arasında pozitif ilişki olduğunda dağılım grafiği nasıl olmalıdır?*
Cevap: Bağımlı y ve bağımsız x değişkenleri arasında negatif (-) ilişki olduğunda bu değişkenlerin dağılım grafikleri aşağıdaki gibi olur:



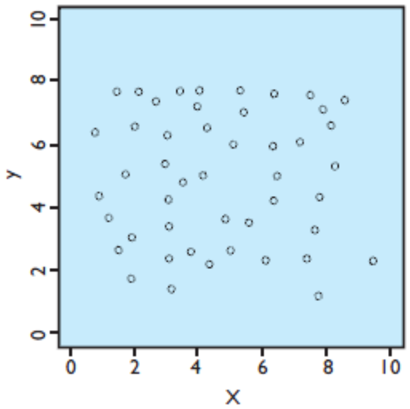
13. *İki değişken arasında rassal bir ilişki olduğunda dağılım grafiği nasıl olmalıdır?*
Cevap: Bağımlı y ve bağımsız x değişkenleri arasında rassal bir ilişki olduğunda bu değişkenlerin dağılım grafikleri aşağıdaki gibi olur:



14. *İki değişken arasında eğrisel bir ilişki olduğunda dağılım grafiği nasıl olmalıdır?*
Cevap: Bağımlı y ve bağımsız x değişkenleri arasında eğrisel bir ilişki olduğunda bu değişkenlerin dağılım grafikleri aşağıdaki gibi olur:



15. *Bağımlı y ve bağımsız x değişkenleri arasında ilişkiyi gözlemlemek için çizilen dağılım grafiği çizilmiş, aşağıdaki şekilde bir grafik elde edilmiştir.*



Buna göre bu iki değişken arasındaki ilişki için korelasyon analizi yapıldığında, değişkenler arasındaki ilişki miktarı nedir?
Cevap: Bağımlı y ve bağımsız x değişkenleri arasında ilişkiyi gösteren dağılım grafiğinde yer alan değişkenler arasındaki ilişkinin rassal olduğu gözlemlenmektedir. Bu nedenle korelasyon analizi yapıldığında aralarında bir ilişki çıkmayacaktır.

16. *İki değişkenler arasındaki ilişkiyi öğrenmek için dağılım grafikleri yeterli midir?*
Cevap: İki değişken arasındaki ilişkinin yalnızca grafikler ile incelenmesi yeterli olmayacaktır. İlişkinin derecesini gösteren istatistiklere ihtiyaç duyulmaktadır.

17. *İki ya da daha fazla, oranlı ve eşit aralıklı ölçeğe uygun şekilde ölçümlenmiş değişkenler arasındaki ilişkinin derecesini belirlemek için hangi korelasyon katsayısı kullanılmalıdır?*
Cevap: İki ya da daha fazla, oranlı ve eşit aralıklı ölçeğe uygun şekilde ölçümlenmiş değişkenler arasındaki ilişkinin derecesini belirlemek için Pearson korelasyon katsayısı kullanılmalıdır**.**

18. *Pearson korelasyon r katsayısının değer aldığı aralıkları yorumlayınız ve şekil üstünde gösteriniz?*
Cevap: Pearson korelasyon katsayısı r, -1 ile +1 arasında değişen değerler almaktadır.

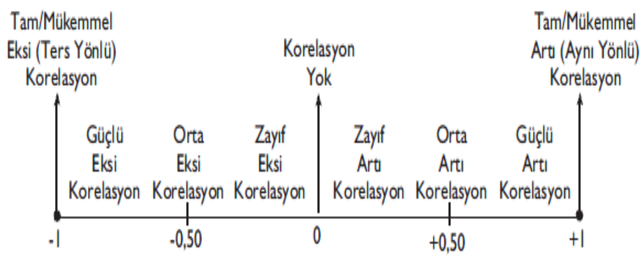
Pearson korelasyon katsayısı r'nin -1 ve +1 değerlerine eşit sonuçlar, mükemmel/tam ilişkinin varlığını gösterir.

Pearson korelasyon katsayısının eksi r değerleri değişkenler arasındaki ters yönlü ilişkiyi gösterirken (biri artarken diğerinin azalması gibi), artı r değerleri değişkenler arasındaki aynı yönlü ilişkinin (biri artarken diğeri de artmaktadır gibi) var olduğunu gösterir.

Eğer iki değişken arasında hiç ilişki yok ise bir başka ifade ile değişkenler bağımsız ise Pearson korelasyon katsayısı 0 (sıfır) değerini alır.

Perason korelasyon katsayısı r, -1 ve +1 değerlerine yaklaştıkça ilişkinin derecesinin arttığı ifade edilirken sıfır değerine yaklaştıkça ilişkinin derecesinin azaldığı/zayıfladığı ifade edilir.

r değeri -0,50 ya da +0,50 etrafında bir değer ise değişkenler arasında orta düzeyli bir ilişkinin varlığı ifade edilir.



19. *Şemsiye satışı yapan bir firma bir yıl içindeki yağmurlu gün sayısı ile şemsiye satışları arasında bir ilişki olup olmadığını araştırmak istemiştir. Buna göre; bir yıl içindeki aylara göre yağmurlu gün sayısını bir değişken ve ilgili aylara göre şemsiye satış miktarını diğer bir değişken olarak tanımlamıştır. Bu amaçla, bir yıl içindeki aylara göre yağmurlu gün sayısı ve şemsiye satış miktarı değişkenlerine ilişkin derlenen veriler aşağıdaki tabloda verilmiştir.*

Aylar	1	2	3	4	5	6	7	8	9	10	11	12
Yağmurlu Gün Sayısı	15	8	18	10	20	10	3	5	6	11	13	13
Şemsiye Satış Miktarı	150	100	200	120	175	70	30	40	70	95	110	290

Cevap: Pearson korelasyon katsayısını hesaplamak için gerekli formül aşağıda verilmiştir:

$$r = \frac{n\sum_{i=1}^{n} x_i y_i - (\sum_{i=1}^{n} x_i)(\sum_{i=1}^{n} y_i)}{\sqrt{n\sum_{i=1}^{n} x_i^2 - (\sum_{i=1}^{n} x_i)^2}\sqrt{n\sum_{i=1}^{n} y_i^2 - (\sum_{i=1}^{n} y_i)^2}}$$

Aylar	Yağmurlu Gün Sayısı (x)	Şemsiye Satış Miktarı (y)	x_i^2	y_i^2	$x_i y_i$
1	15	150	225	22.500	2.250
2	8	100	64	10.000	800
3	18	200	324	40.000	3.600
4	10	120	100	14.400	1.200
5	20	175	400	30.625	3.500
6	10	70	100	4.900	700
7	3	30	9	900	90
8	5	40	25	1.600	200
9	6	70	36	4.900	420
10	11	95	121	9.025	1.045
11	13	110	169	12.100	1.430
12	13	290	169	84.100	3.770
Toplam	132	1.450	1.742	235.050	19.005

Formül yardımıyla tablo verileri kullanılarak gerekli hesaplamalar yapıldıktan sonra; Pearson korelasyon değerine göre iki değişken arasında güçlü ve aynı yönlü korelasyon olduğu söylenebilir.

$$r = \frac{12(19005) - (132)(1450)}{\sqrt{12(1742) - (132)^2}\sqrt{12(235050) - (1450)^2}} = \frac{36660}{\sqrt{11400}\sqrt{718100}} = 0{,}733$$

20. *Araştırmacılar, bağımlı değişkende meydana gelen değişim içerisinde bağımsız değişkenin payının ne olduğunu öğrenmek için kullanılan istatistiğin adı nedir?*
Cevap: Bağımlı değişkende meydana gelen değişim içerisinde bağımsız değişkenin payını gösteren istatistiğe ***belirlilik katsayısı*** adı verilir.

21. *Evrendeki ve örneklemdeki belirlilik katsayısını sembolize eden istatistiğin simgesi genelde ne ile gösterilir?*
Cevap: Evrendeki x ve y değişkenlerinin belirlilik katsayısı ρ^2 ile sembolize edilir. Örneklemdeki x ve y değişkenlerinin belirlilik katsayısı r^2 simgesi ile gösterilir.

22. *Yapılan araştırmada pearson korelasyon katsayısı r=0.733 olarak bulunmuş olduğuna göre araştırmada kullanılan bağımsız değişkenin yüzde kaçını açıklar?*
Cevap: Belirlilik katsayısı bu değerin karesi olan,

r^2 = (0,733) 2 = 0,537 olacaktır.

Yani bağımsız değişken bağımlı değişkenin %53.7'sini açıklar.

23. *Bir araştırmada örneklem sonuçlarına göre belirlenen korelasyon değerinin evren değeri için test edilmesine ne denir?*
Cevap: Bir araştırmada örneklem sonuçlarına göre belirlenen korelasyon değerinin evren değeri için test edilmesine ***korelasyon katsayısı anlamlılık testi*** denir.

24. *Korelasyon katsayısı anlamlılık testi için kurulan hipotez nasıl ifade edilmelidir?*
Cevap: Korelasyon katsayısı anlamlılık testi için kurulan hipotez aşağıdaki gibi ifade edilebilir:

- $H_0 : \rho = 0$
  (x ve y değişkenleri arasında korelasyon yoktur.)
- $H_1 : \rho \neq 0$
  (x ve y değişkenleri arasında korelasyon vardır.)

25. *Korelasyon katsayısı anlamlılığının test edilme adımlarını ifade ediniz?*
Cevap: Korelasyon katsayısı anlamlılığının test edilme adımları şöyle sıralanabilir:

- Adım 1: Hipotezlerin ifade edilmesi,
- Adım 2: Anlamlılık düzeyinin belirlenmesi,
- Adım 3: İstatistiksel test,
- Adım 4: İstatistiksel kararın verilmesi.

BASİT DOĞRUSAL REGRESYON ANALİZİ

26. *Değişkenler arasındaki bu ilişkinin matematiksel bir fonksiyonla ifade edilmesine ne ad verilir?*
Cevap: Değişkenler arasındaki bu ilişkinin matematiksel bir fonksiyonla ifade edilmesi, regresyon analizinin konusunu oluşturur.

27. *Regresyon denklemi, bir bağımlı ve bir bağımsız değişkenden meydana geliyorsa ve değişkenler arasındaki ilişki doğrusal ise yapılan regresyon analizine ne denir?*
Cevap: Bir bağımlı ve bir bağımsız değişkenden meydana gelen doğrusal ilişkiye ***basit doğrusal regresyon analizi*** denir.

28. *Doğrusal regresyon denkleminin tahmini için en çok kullanılan teknik hangisidir?*
Cevap: Doğrusal regresyon denkleminin tahmini için en yaygın kullanılan teknik en küçük kareler tekniğidir.

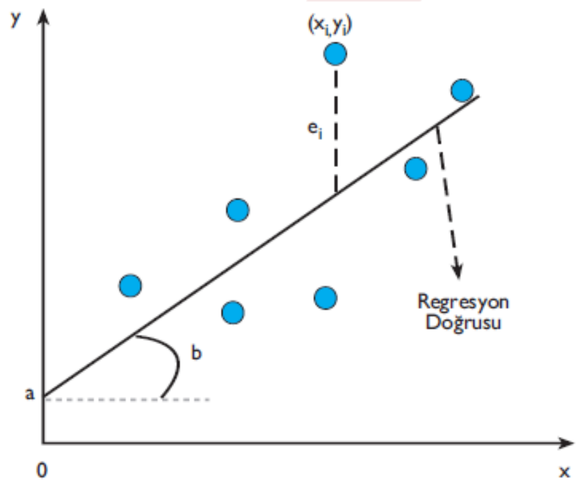
29. *n hacimli örneklem için x ve y değişkenleri arasındaki kuramsal ilişkinin tahmini için kullanılan doğrusal tahmin denklemini ifade edip, denklemde yer alan değişken ve parametrelerin ne anlama geldiğini açıklayınız?*
Cevap: Basit doğrusal regresyon denklemi:

$$y_i = a + bx_i + e_i$$

Bu modelde;

- y_i : bağımlı değişken y'nin i'inci gözlem değerini gösterir.
- x_i : bağımsız değişken x'in i'inci gözlem değerini gösterir.
- e_i : i'inci gözlem için ortalaması sıfır ve tüm gözlemler için sabit σ standart sapmalı normal dağılıma sahip olduğu varsayılan rassal hatayı gösterir.
- a ve b : sırasıyla a ve b 'nin tahmin değerlerini gösterir.
- Benzer şekilde a, x=0 olduğunda y'nin değerini, b ise x'te meydana gelecek birim değişikliğin y'deki oransal etkisini gösterir.

30. *Regresyon modelini grafiksel olarak nasıl ifade edilir?*
Cevap: Tahmin edilen regresyon modelinin grafiksel gösterimi aşağıda verilmiştir:



31. *Regresyon çözümlemesinden elde edilen sonuçların güvenilirliliği hangi varsayımların geçerliliğine bağlıdır?*
Cevap: Eğer varsayımlar geçerli değilse diğer regresyon modelleri ve çözümlemesi yaklaşımlarına başvurulmalıdır. Bu varsayımlar (örneklem için ifade edilmiştir):

- e rassal değişkeni x'in değerlerinden istatistiksel olarak bağımsızdır.
- e rassal değişkeni normal dağılıma sahiptir.
- e rassal değişkeninin sıfır aritmetik ortalamaya sahiptir.
- e_i ve e_j gibi birbirinden farklı iki hatanın istatistiksel olarak bağımsız olmalıdır.
- e_i rassal değişkenleri x_i'lerin tüm değerleri için sabit bir varyansa sahiptir.

32. *Regresyon doğrusunun anlamlılığının test edilmesi için kurulan hipotezi, nedenini ve test edilmesi için kullanılan test hangisidir?*
Cevap: Basit doğrusal regresyonda ilgilenilen iki değişken arasındaki doğrusal ilişki için en küçük kareler tekniği kullanılarak bir regresyon doğrusu tahmini elde edilir. elde edilen bu denklem için kullanılan "bağımlı y_i değerlerini tahmin ederken kullanılan bağımsız x_i değişkeninin değerlerini bilmenin gerçekten faydası var mıdır?" sorusunun cevabının araştırılması gerekmektedir.

Bu araştırma için kurulacak olan "Regresyon doğrusunun anlamlılık testi" hipotezi:

- $H_0 : \beta = 0$ (Regresyon doğrusu anlamlı değildir.)
- $H_1 : \beta \neq 0$ (Regresyon doğrusu anlamlıdır.)

şeklindeki hipotezi t testi kullanılarak test edilmelidir.

Olasılık Kavramı

Olasılık, herhangi bir durumun ya da olayın gerçekleşme şansını ifade eden bir sayı olarak tanımlanır. Sıfır ile bir kapalı aralığında değerler alır.

Olasılık Tanımları

Olasılık tanımlarına ilişkin, nesnel ve öznel olasılık olmak üzere iki yaklaşım bulunur. Nesnel olasılık, klasik olasılık ve deneysel olasılık olmak üzere iki kısımda incelenir. Formülasyonlarda, herhangi bir A olayının gerçekleşme olasılığı P(A) ile belirtilecektir.

Klasik Olasılık: Bir denemenin sonuçlarının eşit olasılıklı olduğu varsayımına dayanır. Klasik bakış açısıyla, bir olayın gerçekleşme olasılığı ilgilenilen sonuçların sayısının, olası tüm sonuçların sayısına bölünmesi yoluyla hesaplanır. Klasik yaklaşıma göre ilgilenilen sonuç sayısı n ile ve olası tüm sonuçların sayısı N ile belirtildiğinde, bir A olayının gerçekleşme olasılığı şu şekilde belirlenir:

$$P(A) = \frac{n}{N}$$

Eğer bir olaylar kümesi bütüne tamamlayan olaylardan oluşuyorsa ve bu olaylar aynı zamanda karşılıklı ayrık olaylar ise olayların gerçekleşme olasılıkları toplamı 1'e eşittir.

Deneysel Olasılık: Bu yaklaşımda bir olayın gerçekleşme olasılığını hesaplamak için, geçmişte benzer olayların gerçekleşme sayısının oranına bakılır. Geçmişte ortaya çıkan olayların sayısı f ile ve toplam gözlem sayısı N ile belirtildiğinde deneysel olasılık yaklaşımıyla bir A olayının gerçekleşme olasılığı şu şekilde hesaplanır:

$$P(A) = \frac{f}{N}$$

Öznel Olasılık: Olasılığın belirlenmesi için herhangi bir geçmiş deneyim ya da bilginin bulunmadığı durumlarda başvurulan yaklaşımdır. Bu yaklaşımda, mevcut görüşler ve eldeki diğer bilgiler değerlendirilerek olaya ilişkin olasılık tahmin edilir.

Olasılık Hesaplama Kuralları

Olasılık hesaplama kuralları, iki veya daha fazla olayın olasılığını hesaplamada kullanılan kurallardır.

Toplama Kuralları

Özel Toplama Kuralı, yalnızca karşılıklı ayrık olaylar için uygulanabilir. Eğer A ve B karşılıklı ayrık olaylar ise özel toplama kuralına göre bu olaylardan birinin veya diğerinin gerçekleşme olasılığı, bu olayların ayrı ayrı gerçekleşme olasılıklarının toplamına eşittir.

$$P(A\ veya\ B) = P(A) + P(B)$$

Tümleyen Kuralı, bir olayın gerçekleşme olasılığının, bu olayın gerçekleşmeme olasılığının 1'den çıkarılması yoluyla belirlendiği kuraldır. A olayının gerçekleşmeme olasılığı $P(\bar{A})$ ise gerçekleşme olasılığı tümleyen kuralına göre şu şekilde hesaplanır:

$$P(A) = 1 - P(\bar{A})$$

Genel Toplama Kuralı, incelenen olayların karşılıklı ayrık olmadığı durumlarda kullanılan olasılık hesaplama kuralıdır. İki ya da daha fazla olayın aynı anda gerçekleşme olasılığı olan *ortak olasılık* kavramı hesaba katılarak genel toplama kuralı şu şekilde hesaplanır:

$$P(A\ veya\ B) = P(A) + P(B) - P(A\ ve\ B)$$

Çarpma Kuralları ve Koşullu Olasılık: Olayların aynı anda gerçekleşme olasılığını hesaplamak için kullanılan kurallardır.

Özel Çarpma Kuralı, A ve B bağımsız olayları için A ve B'nin birlikte gerçekleşme olasılığı, bu iki olayın ayrı ayrı gerçekleşme olasılıklarının çarpımına eşittir. Şu şekilde hesaplanır:

$$P(A\ ve\ B) = P(A).P(B)$$

Burada bahsedilen bağımsızlık kavramı, bir olayın gerçekleşmesinin, bir diğer olayın gerçekleşme olasılığını etkilememesi durumunda bu iki olaya bağımsız olaylardır.

Koşullu Olasılık, B olayının gerçekleştiği bilindiğinde, A olayının gerçekleşme olasılığına koşullu olasılık denir. $P(A|B)$ ifadesi ile gösterilir, B olayının gerçekleşme olasılığı bilindiğinde A olayının koşullu olasılığı olarak ifade edilir.

Genel Çarpma Kuralı, bağımsız olmayan iki olayın ortak olasılığını hesaplamak için kullanılır. Genel çarpma kuralına göre A ve B gibi iki olay için, bu olayların birlikte gerçekleşme olasılığı, A olayının gerçekleşme olasılığı ile A'nın gerçekleştiği bilindiğine göre B'nin koşullu olasılığının çarpımına eşittir. Şu şekilde hesaplanır:

$$P(A\ ve\ B) = P(A).P(B|A)$$

Bayes Teoremi

$\boldsymbol{A_1}$ ve $\boldsymbol{A_2}$ olayları, karşılıklı ayrık ve bütüne tamamlayan olaylar olsun. Bayes teoremine göre, B olayının gerçekleştiği bilindiğine göre $\boldsymbol{A_i}$ 'nin gerçekleşme olasılığı şu şekilde hesaplanır:

$$P(\boldsymbol{A_i}|B) = \frac{P(\boldsymbol{A_i})P(B|\boldsymbol{A_i})}{P(\boldsymbol{A_1})P(B|\boldsymbol{A_1}) + P(\boldsymbol{A_2})P(B|\boldsymbol{A_2})}$$

Sayma Kuralları

Bir denemedeki olası sonuçların sayısı küçük olduğunda bunların sayılması bazen işlemleri kolaylaştırır. Sayma işlemlerini kolaylaştırmak amacıyla saymanın temel ilkesi, permütasyon kuralı ve kombinasyon kuralı olmak üzere üç sayma kuralı kullanılır.

Saymanın Temel İlkesi: Eğer bir işlem m farklı yolla, bir başka işlem de n farklı yola gerçekleşebiliyorsa bu iki işlem birlikte mxn farklı yolla gerçekleşir.

Permütasyon Kuralı: n olası nesnenin tek bir grubundan seçilen r adet nesnenin herhangi bir sıralaması permütasyon kuralı ile hesaplanır. Permütasyonda, sıralanan nesnelerin dizildiği sıra önemlidir. Permütasyon kuralı şu şekilde hesaplanır:

$$_nP_r = \frac{n!}{(n-r)!}$$

Kombinasyon Kuralı: Seçilecek nesnelerin diziliş sırası önemli değilse yapılan her bir seçime bir kombinasyon adı verilir. Kombinasyon kuralına göre, n adet nesneden oluşan bir kümeden r nesnelik kombinasyon sayısı şu şekilde hesaplanır:

$$_nC_r = \frac{n!}{r!\,(n-r)!}$$

Rassal Değişkenler ve Olasılık Dağılımları

Rassal Değişken: Denemeden denemeye farklı değerler lan ve aldığı bu değerleri belli bir olasılıkla alan değişkenlere rassal değişken denir. Rassal değişkenler genellikle büyük harflerle, aldığı değerler ise küçük harflerle belirtilir. Rassal değişkenler, kesikli ve sürekli rassal değişkenler olmak üzere ikiye ayrılır.

Kesikli Rassal Değişken, sonlu ya da sayılabilir sayıda farklı değerler alabilen rassal değişkenlere kesikli rassal değişken adı verilir. Kesikli rassal değişkenler aldıkları değerleri genellikle incelenen olaya konu olan birimlerin sayılması sonucunda alırlar.

Sürekli Rassal Değişken, sayılamayacak ya da sonsuz sayıda olası değerler alabilen ve bir sayı aralığı ya da aralık kümesi üzerinde tanımlanan rassal değişkendir.

Olasılık Dağılımları: Bir denemedeki olası tüm sonuçların ve bu sonuçların her birine ilişkin olasılıkların yer aldığı listeye olasılık dağılımı denir. Olasılık dağılımlarında, herhangi bir denemede ortaya çıkabilecek tüm değerlerin tanım bölgesi verilir. Olasılık dağılımlarının iki temel özelliği vardır:

- Belli bir sonucun olasılığı 0 ile 1 kapalı aralığında değerler alır.
- Tüm karşılıklı ayrık olayların olasılıkları toplamı 1'e eşittir.

Olasılık dağılımları, değişkenin aldığı değere göre kesikli ve sürekli olasılık dağılımları olmak üzere ikiye ayrılır.

Kesikli Rassal Değişkenler İçin Olasılık Dağılımları, Eğer X kesikli bir rassal değişken ise X'in olasılık dağılımı, rassal değişkenin alabileceği tüm olası x değerlerine ilişkin olasılıkların yer aldığı listedir.

Kesikli Bir Olasılık Dağılımın Ortalama, Varyans ve Standart Sapması, Verilerin merkezi eğilimini ve değişkenliğini tanımlamak için ortalama standart sapma veya varyans kullanılır. Bir olasılık dağılımı da benzer biçimde betimlenir. Bir olasılık dağılımının ortalaması μ, varyansı σ^2 ile ve standart sapması σ ile gösterilir.

Kesikli olasılık dağılımları için ortalama şu şekilde hesaplanır:

$$\mu = \sum xP(x)$$

Kesikli olasılık dağılımının varyansı da şu şekilde hesaplanır:

$$\sigma^2 = \sum (x-\mu)^2 P(x)$$

standart sapması ise varyansın karekökü alınarak elde edilir.

Binom Dağılımı, kesikli olasılık dağılımları arasında uygulamalarda en sık rastlanan olasılık dağılımıdır. Binom dağılımının özellikleri şu şekildedir:

- Denemeler, daima aynı koşullarda tekrarlanmalıdır.
- Yapılacak her denemenin sonucunda, var olan karşılıklı ayrık iki sonuçtan yalnızca biri ortaya çıkmalıdır. Bu sonuçlardan biri ilgilenilen sonuç, diğeri ise bunun tümleyeni olan ilgilenilmeyen sonuçtur.
- Rassal değişken, sabit sayıda denemedeki ilgilenilen durumun sayısını belirtir.
- Tek bir denemede ilgilenilen sonucun gerçekleşme olasılığı, tüm denemelerde aynı kalmalıdır.
- Denemeler birbirinden bağımsız yapılmalıdır.

n deneme sayısı, x denemede ilgilenilen sonuçların sayısı olarak tanımlanan rassal değişken, p tek bir denemede ilgilenilen sonucun gerçekleşme olasığı iken binom olasılık dağılımı şu şekilde ifade edilir:

$$P(x) = {}_nC_x p^x (1-p)^{n-x}$$

Sürekli Rassal Değişkenler İçin Olasılık Dağılımları, sürekli rassal değişkenler için deneme sonuçları bir değer aralığı üzerindeki noktalarla belirtilir ve değişkenin aldığı değerler olasılık dağılımları yardımıyla uygun noktalarla ilişkilendirilir.

Normal Dağılım, istatistik teorisinde kullanılan temek olasılık dağılımlarından en önemlisi olan dağılımın günlük yaşamda pek çok uygulaması ile karşılaşılır. Normal olasılık dağılımı şu eşitlik ile gösterilir:

$$P(x) = \frac{1}{\sigma\sqrt{2\pi}} e^{-\frac{1}{2}\left(\frac{X-\mu}{\sigma}\right)^2}$$

burada X sürekli rassal değişkenin aldığı değeri, μ evren ortalamasını, σ evren standart sapmasını gösterir.

Normal dağılımın özellikleri şu şekildedir:

- Şekli çan eğrisi biçimindedir.
- Tek bir tepe noktasına sahiptir.
- Dağılımın aritmetik ortalaması, mod ve medyan değerleri birbirine eşittir ve dağılımın merkezinde yer alırlar.

- Simetrik bir dağılımdır.
- Dağılım eğrisi, merkez değerinden her iki yöne doğru düzgün şekilde azalır.
- Asimptotiktir.
- Normal dağılım eğrisinin altında kalan alan 1'e eşittir.

Normal dağılımın özel bir durumu olan standart normal dağılım, elde edilebilecek tüm normal dağılımlar için olasılık belirlemede kullanılabilmektedir. Aritmetik ortalaması 0 ve standart sapması 1 olan normal dağılıma standart normal dağılım adı verilir.

Herhangi bir normal dağılım, gözlem değerlerinden aritmetik ortalama değeri çıkarıldıktan sonra, bu farkın standart sapmaya bölünmesi yoluyla standart normal dağılıma dönüştürülebilir. Bu işlem sonucu elde edilen değerlere z değerleri ya da standart normal değerler adı verilir. Z değerleri standart değerler olduğu için birimleri yoktur. X gözlem değerleri şu eşitlik yardımıyla z değerime dönüştürülür:

$$z = \frac{X - \mu}{\sigma}$$