

How to calculate p-values

2 tür p-value var: One-sided ve two-sided. Biz two-sided'a odaklanacağız.

Ör-1

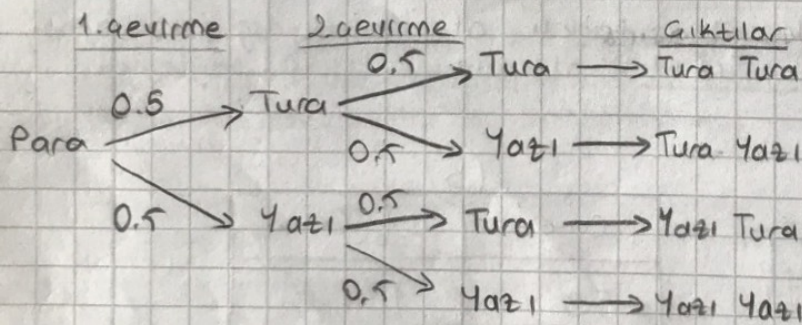
- Diyelim ki madeni bir param var. 2 kez yata-tura attığımdan ikisinde de tura geliyor sonra diyorum ki benim param süper özel çünkü orka orkaya 1kr kez tura geldi.

İstatistiksel bir şekilde hipotezi şu şekilde kurmam daha doğru olur: "Üst üste 2 tura almama rağmen benim paramın normal paramdan farkı yok."

Özel olduğunu düşünmemize rağmen tam tersini **null hypothesis** olarak belirliyoruz ve **küçük p-value** bize onu reddetmemizi söyleyecektir. Eğer null hypothesis'ini reddedersek bileceğiz ki benim param özel, savunduğum gibi.

- p-value'yu hesaplayarak bunu test edeceğiz.

p-value olasılıkların toplanmasıyla belirlenir, o halde ordanda 2 tura gelme olasılığını bularak başlayalım.



2 kez çevirmeden sonra 4 olası çıktı var.

Tüm çıktıların olasılıkları birbirine eşit olduğu için üst üste 2 kez tura gelme olasılığını şu şekilde hesaplayabiliriz.

$$\frac{\text{2 kez tura alma sayısı}}{\text{Toplam çıktı sayısı}} = \frac{1}{4} = 0.25$$

Aynı şekilde üst üste 2 kez yazı gelme olasılığı;

$$\frac{\text{2 kez yazı gelme sayısı}}{\text{Toplam çıktı sayısı}} = \frac{1}{4} = 0.25$$

Bir yazı bir tura gelme olasılığı

$$\frac{\text{Bir yazı ve bir tura gelme sayısı}}{\text{Toplam çıktı sayısı}} = \frac{2}{4} = 0.5$$

Bun olasılığı hesaplarken sırası göz ardı ettik. Çünkü ilk çevirmede yazı veya tura gelmesi ikinci kez çevirmedeki yazı veya tura gelmesini etkilemiyor.

• p-value değeri şu bölümden oluşur.

1) Gözlemlenen olayın gerçekleşme olasılığı

$$\text{2 kez tura gelme olasılığı : } 0.25$$

2) Eşit derecede nadir olan bir olayın gözlemlenme olasılığı

Bizim örneğimizde 2 kez üst üste yazı gelmesi 2 kez üst üste tura gelmesi kadar nadir bir olay olduğu için bunu kabul edebiliriz.

$$\text{2 kez yazı gelme olasılığı : } 0.25$$

3) Daha nadir bir şey gözlemlenme olasılığı

Bizim örneğimizde daha nadir gözlemlenen bir şey yok. Bu yüzden 0.

$$\text{2 tura için p-value} = 0.25 + 0.25 + 0 = 0.5$$

(Null hypothesis)

Bu p-value'yu "Üst aste 2 kez tura almama rağmen benim paramın normal paradan farklı yok" hipotezini test etmek için hesaplamistik.

Tipik olarak, p-value 0.05'ten küçük olduğunda null hypothesisi reddediyoruz.

$0.5 > 0.05$ olduğu için null hypothesisini reddedemiyoruz.

Yani üst aste 2 tura gelmesi bizi madeni paramızın özel olduğunu kanıtlamıyor.

• Neden nadiren veya çok ekstrem görülen olayların olasılığını da ekliyoruz?

→ Nadiren görülen (2) olayları ekliyoruz çünkü her ne kadar 2 kez üst aste tura gelmesi parayı özel yapıyorsa, o kadar nadir görülen bir olayı eklediğimizde o kadar özel görünmeyecektir.

Sevdiğimize bir aralık verip bu aralığın nadiren görülen bir aralık olduğunu söylediklerinde, sevdiğimiz bu aralığın özel bir aralık olduğunu düşünmeyecektir.

Fakat sevdiğimize bu aralığın diğer aralıklar kadar nadir olduğunu söylediklerinde, sevdiğimiz bu aralığın özel olduğunu düşünmeyecektir.

Dolayısıyla nadir görülen olayları eklemek p-value'yu büyütecektir.

→ Çok ekstrem (3) görülen olayları eklememiz de aynı sebepten dolayı gerektiriyor.

Sevdiğimize aralık alıp bu aralığın bu tarafa en nadir aralık olduğunu daha nadiren olmadığını söylediklerinde, sevdiğimiz bu aralığın çok özel olduğunu düşünmeyecektir.

Fakat sevdiğimize bu aralaktan daha nadir görülen aralıklar olduğunu söylediklerinde, sevdiğimiz bu aralığın çok özel olduğunu düşünmeyecektir.

ÖR-2

5 kez madeni parayı attık dgelim, Tura-Tura-Tura-Yazı-Tura geldi, Madeni paraların bu şekilde gelmesi paramı özel yapıyor mı?

Bunu test etmek için p-value'yu hesaplayacağız. Hipotezimi şu şekilde kuruyorum:

4 Tura 1 Yazı almama rağmen benim param normal paradan farklı değildir.

Eğer paramızın özel olup olmadığını öğrenmek istiyorsak Null Hypothesis normal parayı odaklanıyor fakat küçük bir p-value alır ve Null Hypothesisi reddedersek paramızın özel olduğunu bileceğiz.

- 5 kez altıginit paranın 5'si de Tura gelebilir.

TTTTT

→ 1 kez

- 4'ü Tura 1'i Yaşı gelebilir. → 5 kez
- 3'ü Tura 2'si Yaşı gelebilir. → 10 kez
- 2'si Tura 3'ü Yaşı gelebilir. → 10 kez
- 1'i Tura 4'ü Yaşı gelebilir. → 5 kez
- 5'i de Yaşı gelebilir. → 1 kez

+
32 ihtimal var.

4 Tura 1 Yaşı gelmesinin p-value'su

- 1) 4 Tura 1 Yaşı gelmesinin olasılığı

$$\frac{5}{32}$$

- 2) Aynı derecede nadir olan başka bir şey rastgele çıkma olasılığıdır
1 Tura 5 Yaşı gelmesinin olasılığı

$$\frac{5}{32}$$

- 3) Daha nadir veya çok daha ekstrem bir şeyin gerçekleşme olasılığı

5'in de Tura gelmesi

5'in de Yaşı gelmesi

$$> \frac{2}{32}$$

Sonuç olarak 4 Tura 1 Yaşı gelmesinin p-value'su: $\frac{5}{32} + \frac{5}{32} + \frac{2}{32} = 0.375$

Tipik olarak Null hipotezini p-value 0.05'ten küçük olduğu zaman reddettiğimiz için $0.375 > 0.05$ olduğu için null hipotezini reddetmiyoruz.

Birer bir deyişle 4 Tura 1 Yaşı almanın bir paranın ögel olduğu konusunda ikna edemedi.

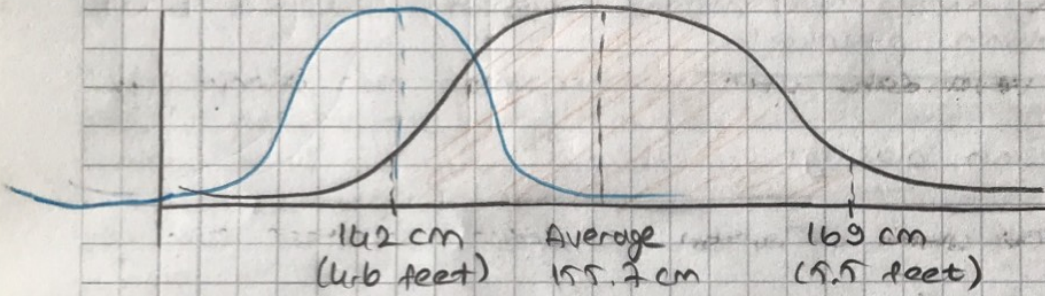
Continuous değerler için p-value

Yüzde tura için tüm çıkışların hepsini kolaylıkla listelemek ve olasılık hesaplamak kolay. Fakat ya insanların ne kadar uzun veya kısa olduğuna ilişkin olasılıkları ve p-value'larını hesaplamak istersek?

Teorik olarak mümkün olan tüm uzunlukları listelemeyi deneyebiliriz

152.4 cm	152.9 cm	153.4 cm	etc...
152.5 cm	153.0 cm	153.5 cm	---
152.6 cm	153.1 cm	153.6 cm	etc...
152.7 cm	153.2 cm	153.7 cm	---
152.8 cm	153.3 cm	153.8 cm	etc...

Pratikte sürekli olan bir şeyin olasılığını ve p-value'larını hesaplamaya kalktığımızda "boy gibi" ağırlıklı istatistiksel dağılım dediğimiz bir şey kullanırız.



Burada 1996 yılında yaşları 15 ile 49 arasında değişen Brezilyalı kadınların boy ölçümlerinin dağılımı verilmiş.

Boyaltı alan Brezilyalı kadınların %95'inin 142 ile 169 cm arasında olduğunu ifade ediyor. Diğer bir deyişle bir Brezilyalı kadının boyunun istatistiksel olarak boyunun 142 ile 169 cm arasında olma olasılığı %95'tir.

Eğri altında kalan alanın %2.5'i 169'dan büyük. Yani bir Brezilyalı kadının boyunun istatistiksel olarak boyunun 169 cm'den büyük olma olasılığı %2.5. Aynı şekilde 142 cm'den küçük olma olasılığı da %2.5.

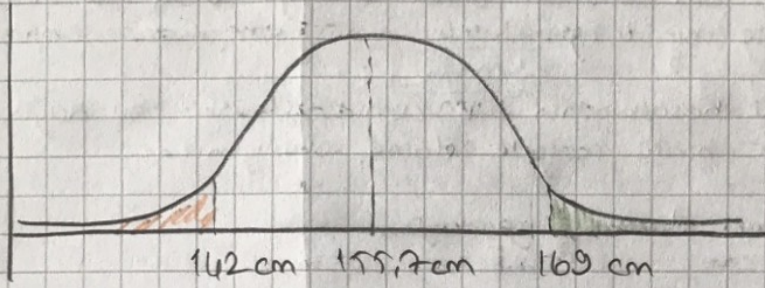
ÖR-1

Bir dağılımın p-value'ları hesaplamak için eğrinin altındaki alan parçalarını toplarız.

Örneğin birinin boyunu 142 cm olarak özetleyelim. Bu boyun ortalama değeri 155.7 olan bu dağılım yüksekliğinde mi yoksa ortalama değeri 142 cm olan başka bir dağılımdan mı geldiğini merak edebiliriz.

Bunu şu: 142 cm olarak özetleyelim bu ölçüm, orijinal dağılımın ortalamasından (155.7) o kadar uzak mı ki, oradan geldiği fikrini reddedelim? Eğer öyleyse başka bir dağılım (başka gibi) veriyi açıklamak için daha iyi bir iş yapıyor olabilir.

"Bu ölçüm (142 cm) sıyah dağılımdan geliyor." hipotezi için p-value 142 cm'den küçük veya eşit kişiler için alanın 0.025'i ile başlar.



Note: Birisi dağılımı kalırken, daha nadir değerlerden ziyade p değerine daha ekstrem değerler eklemekle ilgileniriz.

Ortalamadan (155.7) en az 142 cm kadar uzak olan değerler ekstrem değerler sınıfına girer.

Aynı şekilde 169 cm veya daha uzun olan insanlar için alanın 0.025'i'ni de ekleriz.

"Boyu 142 cm olan birisi sıyah dağılımdan geliyor" hipotezi için

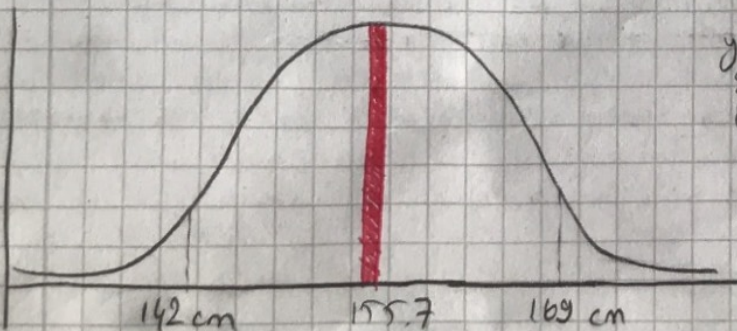
$$p\text{-value} = 0.025 + 0.025 = 0.05 \text{ tir.}$$

Threshold 0.05 olduğu için sıyah dağılımdan da geliyor olabilir, gelmiyor da olabilir. Verimiz yetersiz.

Note: 141 cm boyunda, yani 142 cm'den biraz daha kısa olan birisi olsaydı p-value $0.016 + 0.016 = 0.03$ olabilirdi. $0.03 < 0.05$ olduğu için 141 cm'lik ölçümün sıyah dağılımdan geldiği hipotezini reddedebiliriz. Böylece, bu kadar kısa birisi olmasının oldukça özel olduğu sonucuna varabiliriz. Bu da farklı bir yükseklik dağılımının daha mantıklı olduğunu gösteriyor.

Ör-2

155.4 ile 156 cm boyları arasında birisi olsaydı ne olurdu?



Eğerden zirvesi tam ortalama yükseklikte dir. Bu yüzden sorumuz şu: 155.4 ile 156 cm arasındaki ölçüm, sıyah dağılımın ortalamasından (155.7) o kadar uzak mı ki ortan geldiği fikrini reddedelim?

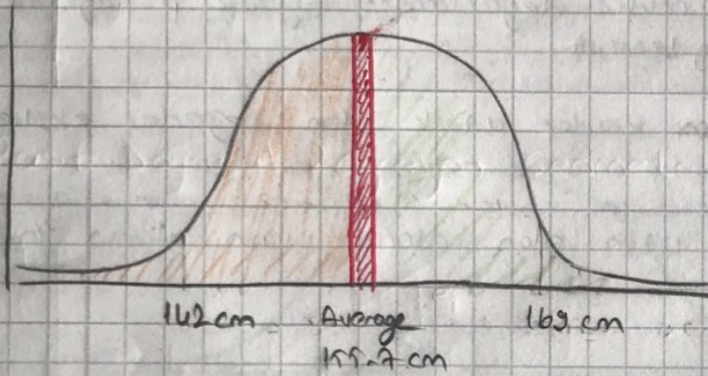
Eğer p-value küçükse, bu başka bir dağılımın yerleri açıklamak için daha iyi bir fit çıkaracağını gösterir.

NOT: Birinin 155.4 ve 156 cm arasında olma olasılığı 0.04. Bu kısmı alan oldukça küçük, neredeyse bir ağız.

Dolayısıyla 0.04, p-value'yi hesaplamamız için kümsür. Çünkü bu dağılıma göre bu değerler aralığındaki birini rastgele ölmeye olasılığıdır.

1. modde

Şimdi daha ekstrem kısımları ölememiz gerekiyor.



Doğru tarafta 155.4'ten daha az olan tüm yükseklikler ortalamadan (155.7) aşağıda. Bu nedenle hepsi ekstrem değerlerdir.

Eğrinin altında kalan alanın %48'i 155.4'ten az uzunluğa sahip. Bu yüzden 0.48'i p-value'ye ekliyoruz.

Aynı şekilde %48'i de 156 cm'den yüksek. Bu yüzden 0.48'i de p-value'ye ekliyoruz.

$$p\text{-value} = 0.04 + 0.48 + 0.48 = 1$$

Yani bu, bu yükseklik dağılımı göz önüne alındığında, olasılığı küçük olan da (0.04) boyu ortalamaya yakın olan birini ölmeye olasılığı bulunmadığı anlamına gelir.