Bölüm 2:Olasılık & Istatistik

dmacpher@coss.fsu.edu

Jump to first page

1. Rakamlar toplamı

dmacpher@coss.fsu.edu

Jump to first page

Rakamlar toplamı

- $\Sigma =$ toplam; X aile geliri gibi değişkendir.
- Daha sonra toplam aile geliri N gözlemlerinin değeri

$$\sum_{i=1}^{N} X_i = X_1 + X_2 + \ldots + X_N$$

dmacpher@coss.fsu.edu

Jump to first page

Rakamlar toplamı

- Sabit zamanlar değişkeni toplamı is değişkenler toplamına eşittir.

$$k\sum_{i=1}^{N} X_i = kX_1 + kX_2 + \ldots + kX_N$$

dmacpher@coss.fsu.edu

Jump to first page

Rakamlar toplamı

- İki değişken gözlemlerinin toplamı onların diğer toplamlarına eşittir

$$\sum_{i=1}^{N}(X_i + Y_i) = \sum_{i=1}^{N} X_i + \sum_{i=1}^{N} Y_i$$

dmacpher@coss.fsu.edu

Jump to first page

Rakamlar toplamı

- N üzerindeki sabit gözlemlerin toplamı sabit ürün ve N e eşittir.:

$$\sum_{i=1}^{N} k = kN$$

dmacpher@coss.fsu.edu

Jump to first page

2. Tanımlar

dmacpher@coss.fsu.edu

Jump to first page

Rastgele Deney

- Rastgele deney: a process leading to at least two possible outcomes with uncertainty as to which will occur.
  - Sonuçlar çözülebilir
    - örneğin - ölmek, madeni para atmek, desteden kart çekmek sastgele deneydir.

dmacpher@coss.fsu.edu

Jump to first page

Örnek alan

- Bir deneyin tüm olasılık sonuçlarının kurulumu is örnek kütle olarak tanımlanır.
- Örnekler :
  - madeni para atmak S={H,T}.
  - ölüm S={1,2,3,4,5,6}
  - 3 sonuçlu madeni para atmak S={HHH, HHT, HTH, THH, HTT, THT, TTH, TTT}

dmacpher@coss.fsu.edu

Jump to first page

Örnek noktalar

- Örnek noktaörnek alanın sonuçlarının herbiridir.
  - Örnekler :
    - Y veya T
    - 1,2,3,4,5 veya 6
    - YYY

dmacpher@coss.fsu.edu

Jump to first page

Sonuçlar

- An event is a subset of the sample space.
  - Örnekler :
    - Sonuç A iki madeni para atıldığında yazı gelmesi. Sonuç TT sonuç A ya aittir.sonuç a 3 madeni para atıldığında önemli etkiye sahiptir.A={HHH, HHT, HTH, THH}

dmacpher@coss.fsu.edu

Jump to first page

Ortak Özel Olaylar

- Sonuçlar özeldir if thebir olayın sonucu diğer olayın sonucunun önüne geçer.
  - A 2 yazıdır ve b 2 başlıdır.ve bunlar ortaktır.
  - Eğer sonuç A önüne geçmezse ve maça ası sonuç B ise bunlar ortaktır.

dmacpher@coss.fsu.edu

Jump to first page

Benzer olayların sonuçları

- Bir olay diğer olaya benziyorsa sonuçlar eşittir.
  - Örnek :
    - 1/6 olasılıkla ölüme atlmaktır

dmacpher@coss.fsu.edu

Jump to first page

Ortak ayrıntılı sonuçlar

- Sonuçlar ortaktır tüm olasılıkların sonuçlarıdır.Example:
    - HH, HT, TH, TT (2 tura, 2 yazı,1 yazı 1 tura)

dmacpher@coss.fsu.edu

Jump to first page

3. Olasılık tanımları

dmacpher@coss.fsu.edu

Jump to first page

Klasik tanımlar

- n olasılık sonuçlarıyla deney farzedelim , ve bu sonuçlar are benzer ve ortak sonuçludur.
  - Örnek kütlede a sonucu kuralım.
  - P(A) A'nın olasılığıdır.
  - M sonuçları olumludur, daha sonraP(A) m/n dır

dmacpher@coss.fsu.edu

Jump to first page

Klasik tanım örneği

- 2 madeni paramümkün sonuçlar (=n) are TT, TH, HT, HH.
- 2 madeni paradan A sonucu kuralım.
  - m=1.
    - bundan dolayı P(A) = 1/4.

dmacpher@coss.fsu.edu

Jump to first page

Kalasik örnek tanımları

- 2 zar atalım, n = 36 mümkün sonuçlar
  - (12,13,14,15,16;21,22,23,24,25,26 ; 31,32,33,34,35,36;41,42,43,44, 45,46; 51,52,53,54,55,56; 61,62,63,64,65,66.)
- Farz edelim A sonucu 7 gösterimdedir.
  - m=6 ve bundan dolayı P(A) = 6/36 veya 1/6

dmacpher@coss.fsu.edu

Jump to first page

Amprik tanımlar

- Klasik tanımda problem - belki deneyin tüm sonuçları eşit olmayabilir.
  - Sonuçların numaraları sonlu değildir..
- Bundan dolayı olasılık tanımı görecelidir

dmacpher@coss.fsu.edu

Jump to first page

Göreceli frekans örneği

- Burada yağış miktarı örneği

Inches	Days	Relative Frequency
<4	10	0.1
5-9	20	0.2
10-14	45	0.45
15-19	20	0.2
>20	5	0.05
Total	100	1

İnc birleştirilmiş sıklık grublarının rastgele değeridir

dmacpher@coss.fsu.edu

Jump to first page

Amprik olasılık tanımları

- Sonsuz sıklıklar olayların numaralarınınn basitidir.
- Göreceli sıklıklar tüm olaylar tarafından bölünür.
- Bu m/n olasılığına benzer.
  - Göreceli sıklıkları olasılıklar gibi ele alabilirmiyiz , n genişse.
  - Bu tanımda, sonuçlar benzer sonuçlara ve ortak ayrıntılara ihtiyacı yoktur.

dmacpher@coss.fsu.edu

Jump to first page

4. Probability Properties

dmacpher@coss.fsu.edu

Jump to first page

Bir sonucun olasılığı

- Bir sonucun olasılığı 0 ve 1 arasındadır.
  - $0 \leq P(A) \leq 1$
  - Eğer P(A)=0 sonuç ortaya çıkmayacaktır.
  - Eğer P(A)=1 sonuç ortaya çıkacaktır.
  - Sonucun olasılığı bu uçlar arasındadır.

dmacpher@coss.fsu.edu

Jump to first page

Ortak ayrıntılı sonuçlar

- Eğer A, B ve C ortaksa, onlardan birinin olasılığıdiğer olasılıkların toplamına eşit olacaktır.
  - P(A+B+C)= P(A) + P(B) + P(C)

dmacpher@coss.fsu.edu

Jump to first page

Ortak özellikler ve ayrıntılı sonuçlar

- Eğer A, B, ve C ortak özellikli ve ortak ayrıntli sonuçlarsa daha sonra bireysel olasılıkların toplamı 1dir.
  - P(A+B+C)=P(A) + P(B) + P(C)= 1

dmacpher@coss.fsu.edu

Jump to first page

Ortak özellikler ve ayrıntılı sonuçlar

- Throw die ve bundan dolayı 1,2,3,4,5,6.dır
  - Bu sonuçlar ortak özellikli ve ayrıntılı sonuçlardir.
    - Birinin olasılığı: 1/6
    - Bundan dolayı P (1,2 veya 3) = P(1) + P(2) + P(3) = 1/6 + 1/6 + 1/6=1/2
    - Ve P(1+2+3+4+5+6) = P(1)+ P(2)....=1

dmacpher@coss.fsu.edu

Jump to first page

İstatiksel bağımsız sonuçlar

- A,B ve C bağımsız sonuçlardır eğer onların olasılığı onların olasılığının ürünüdür.
  - Eğer P(ABC) = P(A)P(B)P(C)
    - P(ABC) birleşik olasılık olaral tanımlanır
    - P(A) ve P(B) ve P(C)marjinal şartsız ve özel olasılık olarak tanımlanır.

dmacpher@coss.fsu.edu

Jump to first page

İstatiksel bağımsız sonuç örneği

- 2 zar attığında 6yı sağlamanın olasılığı sallamadır
  - Bir altı olasılığı A
  - Bir altı olasılığı B
  - Bundan dolayı P(AB) = P(A) P(B) = (1/6)(1/6)= 1/36

dmacpher@coss.fsu.edu

Jump to first page

Şartlı olasılık

- farzedelim sonuç B nin olasılığını istiyoruz bilinen A meydana geldimi?
  - Bu şartlı olasılıktır
    - P(B\A) = P(AB)/P(A)

dmacpher@coss.fsu.edu

Jump to first page

Şartlı olasılık örneği

- Bir desteden kart seçileceğini farz edelim.sinek kralı çıkma olasılığı nedir?
  - Sonuç A kral sonuç B ise sinek
  - P(B/A) = P(AB)/P(A)
    - P(B/A) = 1/52 / 4/52 = 1/4

dmacpher@coss.fsu.edu

Jump to first page

5. Olasılık sıklık fonksiyonu

dmacpher@coss.fsu.edu

Jump to first page

Olasılık dağılım fonksiyonu

- Olasılık yoğunluk fonksiyonu (PDF) rastgele değişkenler olasılığını belirler.
- Rastgele değişkenler için X, pdf f(X) anlamına gelir.
- pdf x değerlerinin karşısındakı olasılıkların dağılımını gösterir.
- Pdf nin toplamı 1 dir.

dmacpher@coss.fsu.edu

Jump to first page

Olasılık dağılımı fonksiyon örneği

- 2 madeni para atalım ve 4 sonuç elde ederiz.YY TT TY YT
- şimdi, 2 madeni para atıldığında rastgele değişkenlerin tanımlayalım

X	f(X)
0	.25
1	.50
2	.25

dmacpher@coss.fsu.edu

Jump to first page

Discrete Probability Distribution Function

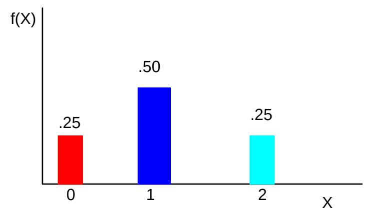
- The discrete PDF of X is $f(x) = P(X=x_i)$ for i=1,2,...n and $f(x) = 0$ for $X \neq x_i$ where $P(X= x_i)$ is probability that the discrete random variable takes the value of x_i.
- Bu rastgele değişkenler büyük harfalerle tanımlanmıştır X, Y etc. (and diğerleri küçük harf, x, y, etc.).

dmacpher@coss.fsu.edu

Jump to first page

Discrete Probability Distribution Function

- Below is the discrete PDF for the two tails example

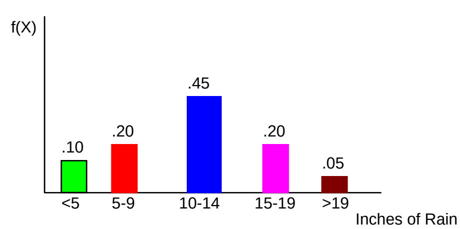


dmacpher@coss.fsu.edu

Jump to first page

PDF of a Continuous Variable

- In this case,aralığın üstündeki rast gele değişkenlerin olasılığını ölçeriz.
  - Sonuçların sınırsız numaraları



dmacpher@coss.fsu.edu

Jump to first page

Kümülatif yoğunluk fonksiyonu (CDF)

- CDF $F(X)=P(X \leq x)$ dir veya rastgele değişkenlerin olasılığı X küçük x ten daha büyük değer taşır.
  - Formül , $F(X)=\Sigma^{x} f(x)$

dmacpher@coss.fsu.edu

Jump to first page

Multivariate PDF

- Eğer sonuçlar bir rastgele değişkenden daha fazla yayılmışsa, daha sonra farklı olasılık dağılım fonksiyonlarına sahip oluruz.
- Bivariate PDF iki rastgele değişken yayılır.
- formul, the discrete joint PDF is $f(X,Y) = P(X=x \text{ ve } Y=y)$ ve $f(X,Y) = 0$ when $X \neq x$ and $Y \neq y$

dmacpher@coss.fsu.edu

Jump to first page

Farklı PDF fonksiyon örneği

- X in birleşik değeri (kolej eğitimi) veY (aile geliri).
  - X = 1eğer head kolej eğitimi ise ; 0 – aksi takdirde
  - Y = $5,000 $10,000 or $15,000

		X	
Y	0	1	Total
$5,000	24	0	24
$10,000	12	12	24
$15,000	16	32	48
Total	52	44	96

dmacpher@coss.fsu.edu

Jump to first page

farklı PDF örneği

- Görecelisıklıkların dönüşümü (olasılıklar).

		X	
Y	0	1	Total
$5,000	0.250	0.000	0.250
$10,000	0.125	0.125	0.250
$15,000	0.167	0.333	0.500
Total	0.542	0.458	1.000

This is a discrete bivariate PDF. Each cell is a joint probability of college education and income.

dmacpher@coss.fsu.edu

Jump to first page

Marjinal PDF

- f(X) ve f(Y) aremarjinal PDF
- Xin marjinal PDF x olasılığı y nin mutlak değeridir.

dmacpher@coss.fsu.edu

Jump to first page

Marjinal PDF

Y	f(Y)	X	f (X)
$5,000	0.250	0	0.542
$10,000	0.250	1	0.458
$15,000	0.500	Total	1.000
Total	1.000		

dmacpher@coss.fsu.edu

Jump to first page

Şartlı PDF

- Y nin şartlı PDF verilmiş x değerinin y olasılığıdır
- f(Y|X) = P(Y=y | X=x) where f(Y | X) y nin şartlı PDF dir..
  - Gibi hesaplanır: f(Y | X)= f(X,Y)/f(X)

dmacpher@coss.fsu.edu

Jump to first page

Şartlı PDF örneği

- Gelir 10000$ olduğunda kolej eğitimi ne olur?
  - f(Y=$10,000| X=1)
  - = f(Y=$10,000 and X=1)/f(X=1)
  - = .125 / .458 = .273

dmacpher@coss.fsu.edu

Jump to first page

İstatiksel bağımsızlık

- X ve y değişkenleri istatiksel bağımsızdır eğer onların marjinal PDF leridir.
  - f(X,Y) = f(X)* f(Y)

dmacpher@coss.fsu.edu

Jump to first page

İstatiksel bağımsızlık örneği

- Let X=H ilk atılan T if tail and let Y=H if head on second coin and T if tail

	X		
Y	H	T	Total
H	0.25	0.25	0.5
T	0.25	0.25	0.5
Total	0.5	0.5	1

f(X=H, Y=H) =.25
f(X=H) = .5
f(Y=H) = .5
So f(X=H, Y=H) = f(X=H) * f(Y=H)

dmacpher@coss.fsu.edu

Jump to first page

İstatiksel bağımsızlık örneği

- Gelir ve eğitim bağımsız değişken değildir.
  - f(Y=$10,000, X= 1) .125 dır.
  - Aynı değildir.: f(Y=$10,000) = .25 * f(X=1) = .458

dmacpher@coss.fsu.edu

Jump to first page

6. Beklenen değer

dmacpher@coss.fsu.edu

Jump to first page

PDFİN MOMENTİ

- Dağılımın birinci momentibeklenen değer veya anlmdır.
- İkinci moment yaryanstır.
- Beklenen değer veya popilasyon anlam değeri:

$$E(X) = \sum_{i=1}^{N} X_i f(X_i)$$

dmacpher@coss.fsu.edu

Jump to first page

Beklenen değer örneği

X	f(X)	Xf(X)
-2	0.27	-0.54
0	0.12	0
2	0.26	0.52
3	0.35	1.05
E(X) = 1.03		

dmacpher@coss.fsu.edu

Jump to first page

Beklenen değer özelliği

- 1. sabitin beklenen değeri sabittir.
  - E(b)=b where b is a constant
    - Ör: E(5)=5
- 2. iki rastgele değişkenlerin toplamının bekleneni rastgele değişkenlerin beklenenleri toplamına eşittir.
  - $E(X + Y) = E(X) + E(Y)$

dmacpher@coss.fsu.edu

Jump to first page

Beklenen değer özelliği

- 3. iki farklı değişkenin beklenen değeri rastgele değişkenlerin beklenen değerlerine eşit değildir.$E(X/Y) \neq E(X)/E(Y)$
    - Ex: ücret anlamı ≠ kazanç/saat .

dmacpher@coss.fsu.edu

Jump to first page

Beklenen değer özellikleri

- 4. iki rastgele değişken ürünün beklenen değeri ratgele değişkenlerin beklenen değerlerinin ürününe eşit değildie genellikle.
  - E(XY) ≠ E(X)E(Y)
    - Bununlaberaber eğer X ve Y bağımsız rastgele değişkenlersee doğrudur.

dmacpher@coss.fsu.edu

Jump to first page

Beklenen değer özelliği

- 5.zaman zartlı rastgele deeğişkenlerin bekleneni,rastgeelee deeğişkeenleerin şarrtlı zamanlarına eşittir.
- E(bX) = bE(X) yerde b şartlıdır.

dmacpher@coss.fsu.edu

Jump to first page

7. Varyans

dmacpher@coss.fsu.edu
Jump to first page

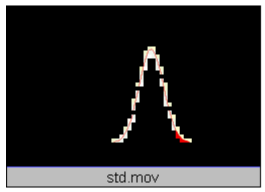
Varyans

- varyans anlamın trafında dağılım gösterir.
  - $var(X)=\sigma^2 = E[(X-\mu)]^2$, yerde μ = E(X) veya beklenen değer
  - varyans x ve beklenen değeriarasında karesinin beklenen değeridir.
  - Standart sappma varyansın karesidir..

dmacpher@coss.fsu.edu
Jump to first page

Varyans

- X anlamın etrafında dağılmışsa X geniştir,



dmacpher@coss.fsu.edu

Jump to first page

Varyans hesaplanması

- Varyans hesaplanması

$$Var(X) = \sum_{i=1}^{N} f(X_i)[X_i - E(X)]^2$$

or

$$Var(X) = E[(X - E(X)]^2$$

dmacpher@coss.fsu.edu

Jump to first page

Variance Example

- Varyans hesaplama örnekleri

X	f(X)	$[X-E(X)]^2$	$[X-E(X)]^2 f(X)$	X^2	$E(X^2)$
-2	0.27	9.18	2.48	4	1.08
0	0.12	1.06	0.13	0	0
2	0.26	0.94	0.24	4	1.04
3	0.35	3.88	1.36	9	3.15
			4.21		5.27
E(X) = 1.03 and Var (X) = 421					
or Var (X) = 5.27 - $(1.03)^2$ = 4.21					

dmacpher@coss.fsu.edu

Jump to first page

Varyansın özellikleri

- 1. şartlı varyans sıfırdır.
- 2. x ve Y iki Bağmsız değişkensex ve y are 2 independent random variables, tonları toplamı veya farkların varyansı özel varyanslarının toplamına eşittir.
  - var (X+Y) = var(X) + var(Y)
  - var (X-Y) = var(X) + var(Y)

dmacpher@coss.fsu.edu

Jump to first page

Varyansın özelliklerei

- Değişkenlere şart eklemek onun varyansını değiştirmez.
  - B şartlıysa var (X + b) = var (X)
- 4. b şartlıysab zaman değişkenlerinin varyansıdeğişkenlerin varyansının zaman şartının karesine eşittir., var(bX)= b^2 var(X)

dmacpher@coss.fsu.edu

Jump to first page

Varyans özellikleri

- 5.a ve b sartlıysa var (aX + B)= a^2 var(X)
- 6. , X ve Y bağımsız rastgele değişkenlerse a ve b şartlıdır.
  - var (aX + bY) = a^2 var(X)+b^2 var(Y)

dmacpher@coss.fsu.edu

Jump to first page

8 kovaryans

dmacpher@coss.fsu.edu

Jump to first page

Kovaryans

- Kovaryans değişkenın birlikte hareketlerini ölçer.
- Farz edelim 2 rastgele değişkene sahibiz , kovaryans:
  - Cov(X,Y)= E[(X-E(X)) (Y-E(Y))]
  - or Cov(X,Y) = E(XY) $-\mu_X \mu_Y$

dmacpher@coss.fsu.edu

Jump to first page

Kovaryans hesaplanması

- $Cov(X,Y) = \Sigma_X \Sigma_Y (X-\mu_X)(Y-\mu_Y) f(X,Y)$
- $Cov(X,Y) = \Sigma_X \Sigma_Y XY\ f(X,Y) - \mu_X \mu_Y$
  - Double summation çünkü değişkenlerin sıralamasının karşısındaki her iki değişkeni topluyoruz. Sürekli rastgele değişkenler kullanımı için gereklidir..
- Eğer 2 değişken beraber hareket ediyorsa,kovaryans pozıtıf.farklı hareket ediyorsa negatiftir.

dmacpher@coss.fsu.edu

Jump to first page

Kovaryans örneği

- Ortak dağılım farz edelim

		X			
Y	-2	0	2	3	f(Y)
3	0.27	0.08	0.16	0.00	0.51
6	0.00	0.04	0.10	0.35	0.49
f(X)	0.27	0.12	0.26	0.35	1.00

dmacpher@coss.fsu.edu

Jump to first page

Kovaryans örneği

X	f(X)	Xf(X)	Y	f(Y)	Yf(Y)	X^2	$E(X^2)$	Y^2	$E(Y^2)$
-2	0.27	-0.54	3	0.51	1.53	4	1.08	9	4.59
0	0.12	0	6	0.49	2.94	0	0	36	17.64
2	0.26	0.52				4	1.04		22.23
3	0.35	1.05				9	3.15		
							5.27		
E(X) = 1.03 and E(Y) = 4.47									
var(X) = 5.27-$(1.03)^2$ = 4.21									
var(Y) = 22.23-$(4.47)^2$ = 2.25									

dmacpher@coss.fsu.edu

Jump to first page

Kovaryans örneği

- Cov (X,Y) = $\Sigma_X \Sigma_Y XY\ f(X,Y) - \mu_X\ \mu_Y$
- $\Sigma_X \Sigma_Y XY\ f(X,Y)$ = (-2)(3)(0.27) +(0)(3)(0.08)+(2)(3)(0.16)+(3)(3)(0) +(-2)(6)(0) + (0)(6)(0.04) + (2)(6)(0.10)+(3)(6)(0.35)
  - = -1.62 + 0.96 + 1.2 + 6.3 = 6.84
- Cov (X,Y) = 6.84 - (1.03)(4.47) = 2.24

dmacpher@coss.fsu.edu

Jump to first page

Kovaryansın özellikleri

- X ve Y bağımsız rastgele değişkenler isekovaryansları sıfırdır.
  - If 2 r.v.s are independent, then:
    - $E(XY) = E(X)E(Y) = \mu_X \mu_Y$
    - kovaryansta yerine koyalım
- cov (a+bX, c+dY) = bd cov (X,Y) yerde a, b, c,d şartlıdır.

dmacpher@coss.fsu.edu

Jump to first page

Kovaryansın özellikleri

- Değişkenlerin kovaryansı onun basit varyansıdır.

dmacpher@coss.fsu.edu

Jump to first page

Korelasyon katsayısı

- eğer değişkenler beraber hareket ediyorsa kovaryans tanımlar.
- Korelasyon katsayısı değişkenin ne kadar güçlü ilişkisini gösterir. $\rho = cov(X,Y)/\sigma_X \sigma_Y$
- Korelasyon katsayısı örneği:
  - $\sigma_X = 2.05$, $\sigma_Y = 1.50$, cov (X,Y) = 2.24.
    - $\rho = 2.24/(2.05)(1.5) = 0.73$

dmacpher@coss.fsu.edu

Jump to first page

Korelasyon katsayısının özellikleri

- Korelasyon katsayısı kovaryans ile aynı işarete sahip.
- Korelasyon katsayısı -1 ve +1 arasındadır
  - kovaryans ölçülmüş X ve Y bölümlerine bağlıdır, fakat korelasyon katsayısı değildir.
  - eğer $\rho = 1$ ise anlamı iki değişken tam pozıtıf korelasyona sahiptir.

dmacpher@coss.fsu.edu

Jump to first page

Korelasyon varyansının değişkenleri

- 2 değişken bağımsızsa korelasyon vardır
  - var(X+Y)=var(X)+var(Y)+2cov(X,Y)
  - var(X-Y) =var(X)+var(Y)- 2cov(X,Y)
- bununla beraber bağımsızlarsa, cov (X,Y) = 0
  - Bundan dolayı var(X+Y) = var(X) + var(Y)

dmacpher@coss.fsu.edu

Jump to first page

Şartlı beklentiler

- Şartlı beklentiler: x in beklene değeri ne ise,y değerleri verilir. $E(X|Y=y) = \Sigma_X X\ f(X|Y=y)$
  - f(X|Y=y) X PDF leri şartlıdır ve x değerlerini toplarız.

dmacpher@coss.fsu.edu

Jump to first page

Çarpıklık and Kurtosis

- Çarpıklık pdf in asimetrisini ölçer.
- $= [E(X- \mu_X)^2]^3 / [E(X- \mu_X)^3]^2$
  - Eğer S >0, doğru çarpıklık and vice versa
- Kurtosis is a measure of the tallness or flatness of the pdf.
  - $K = E(X- \mu_X)^4 / [E(X- \mu_X)^2]^2$
  - K>3: tall, K<3: flat

dmacpher@coss.fsu.edu

Jump to first page

9. Toplamdan örneğe

dmacpher@coss.fsu.edu

Jump to first page

Toplamdan örneğe

- Beklenen değerleri hesaplamak, pdf in varyansı,örnek kütlenin tümüne veya toplam kütleye ihtiyacımız var.
- pratikte toplamdan temsil eden örnek kullanırız.
  - Bu bize anlam varyans etc tahminini verir.

dmacpher@coss.fsu.edu

Jump to first page

Örnek anlam

- Örnek anlam E(X) in tamın edicisidir.

$$\overline{X} = \sum_{i=1}^{n} X_i / n$$

dmacpher@coss.fsu.edu

Jump to first page

Örnek varyans

- örnek varyans σ^2ın toplam varyanstan tahmin edicisidir.

$$S_x^2 = \sum_{i=1}^{n} \frac{(X_i - \overline{X})^2}{n-1}$$

n-1 serbestlik derecesidir.
N e bölersekdaha sonra n-1, we get a biased estimate since only n-1 unconstrained observations exist.

dmacpher@coss.fsu.edu

Jump to first page

Örnek kovaryans

- Örnek kovaryans toplam kovaryansın tahmin edicisidir.

$$\text{Sample Cov(X,Y)} = \frac{\sum (X_i - \overline{X})(Y_i - \overline{Y})}{n-1}$$

dmacpher@coss.fsu.edu

Jump to first page

Örnek korelasyon katsayısı

- Örnek korelasyon katsayısı toplam korelasyon katsayısının tahmin edicisidir

$$r = \frac{\sum_{i=1}^{n}(X_i - \overline{X})(Y_i - \overline{Y})/(n-1)}{S_X S_Y}$$

dmacpher@coss.fsu.edu

Jump to first page

10. Normal dağılım

dmacpher@coss.fsu.edu

Jump to first page

Normal dağılım

- simetriktir, bell-shaped distribution.
- Normal dağılmış sürekli rastgele değişken X in pdf:

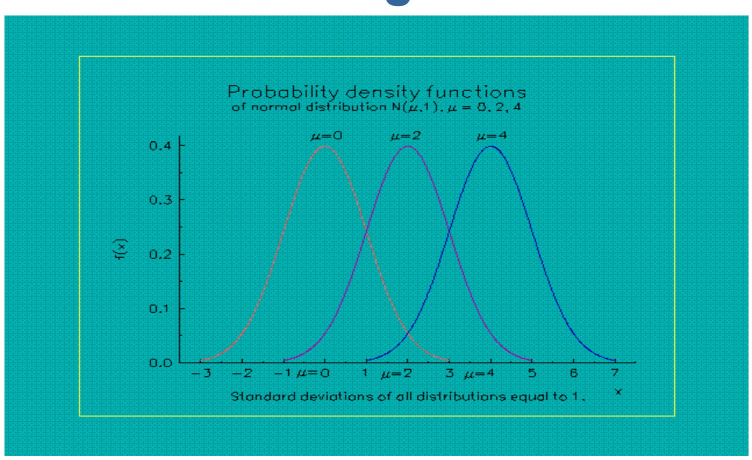
$$f(X) = \frac{1}{\sigma\sqrt{2\pi}} \exp\left[-\frac{1}{2}\left(\frac{X-\mu}{\sigma}\right)\right]$$

Anlam ve varyans tarafından tanımlanmıştır: $X \sim N(\mu,\sigma^2)$

dmacpher@coss.fsu.edu

Jump to first page

Normal dağılım



dmacpher@coss.fsu.edu

Jump to first page

olasılıklar

- Normal dağılımın olasılıkları:
  - olasılık(μ-σ<X< μ+σ) $\cong$.68
  - olasılık (μ-2σ<X< μ+2σ) $\cong$.95
  - olasılık(μ-3σ<X< μ+3σ) $\cong$.997

dmacpher@coss.fsu.edu

Jump to first page

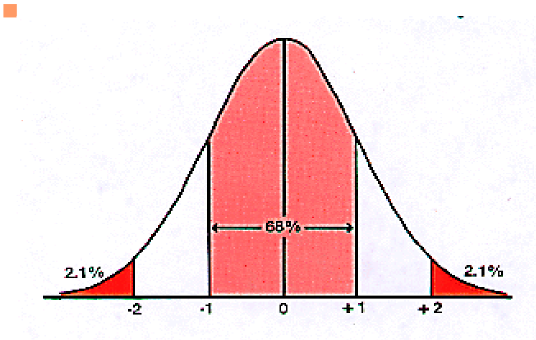
Standart normal dağılım

- μ ve σ parametreleri bilinirse,x aralığının olasılıklarını bulabiliriz.
  - Normal dağılmış x değişkenlerini m ile z içinde değiştirmek.
  - $Z=(X - \mu_x)/ \sigma_x$
    - Standartlaştırılmış değişkenlerThis standardized variable, Z, 1 varyansa ve 0 anlama sahiptir.
    - Thus Z~N(0,1) - a standart normal

dmacpher@coss.fsu.edu

Jump to first page

Standart normal dağılım



dmacpher@coss.fsu.edu

Jump to first page

Z-sonuç tablosu (Tablo A-1)

Z	Area
0.0	0.0000
0.5	0.1915
1.0	0.3413
1.5	0.4332
2.0	0.4772
2.5	0.4938
3.0	0.4987

dmacpher@coss.fsu.edu

Jump to first page

Standart normal dağılım örneği

- X~N(8,4) farzedelim. X 4 ve 12 arsında değerse olasılık ne?
  - $Z1=(X1-\mu_x)/\ \sigma_x=(4-8)/2 = -2$
  - $Z2=(X2-\mu_x)/\ \sigma_x=(12-8)/2 = 2$
- Tabloya bak: $p(0\leq Z \leq 2)= 0.4772$
- Simetrik dağılım: $p(-2\leq Z \leq 0)=$ 0.4772
- Bundan dolayı olasılık = 0.4772+0.4772= 0.9544

dmacpher@coss.fsu.edu

Jump to first page

Standart normal dağılım örneği

- SAT sonuçlar ~N (500,10000) farz edelim herhangi bir sonucun 750 üzerinde olma olasılığı ne?
- $Z=(X1-\mu_x)/\ \sigma_x=(750-500)/100 =2.5$
  - Pr(0<Z<2.5)=0.4938
  - So Pr (Z>2.5) = 0.5-0.4938 = .0062 or .6 percent
- Note that if σ_x is greater (e.g. 200)
  - Z sonuçları daha az & olsılık artar.

dmacpher@coss.fsu.edu

Jump to first page

Doğrusal birleşme

- 2 normal dağılmış rastgele değişkenin doğrusal birleşimi onun normal dağılımıdır.

If $X \sim N(\mu_X, \sigma_X^2)$ and $Y \sim N(\mu_Y, \sigma_Y^2)$ and they are independent

Consider this linear combination of X and Y:

$$W = aX + bY$$

Then it can be shown that:

$$W \sim N[(a\mu_X + b\mu_Y), (a^2\sigma_X^2 + b^2\sigma_Y^2)]$$

This can extend to more than 2 random variables

dmacpher@coss.fsu.edu

Jump to first page

11. Örnek anlamın dağılımı

dmacpher@coss.fsu.edu

Jump to first page

Toplam ve örnek anlam

- Toplam anlam tüm dağılımın beklenen değeridir.
- Örnek anlam verilen gözlemlerin anlamıdır.
  - Farklı örnekler farklı örnek anlamlarına benzemektedir.Different samples

dmacpher@coss.fsu.edu

Jump to first page

Sampling of Means

- Assume a random sample of $X_1, X_2, \ldots, X_n$ from a population PDF with μ and σ^2.
  - Random sample - the r.v.s must be independent and identically distributed (i.i.d)
    - PDF of each X_i is ~ $N(\mu, \sigma^2)$

$$\text{Let } \bar{X} = X_i / n$$

$$\text{then } \bar{X} \sim (\mu, \sigma^2 / n)$$

dmacpher@coss.fsu.edu

Jump to first page

Sampling of Means

- Konu ile ilgili Z sonuçları:

$$Z = \frac{\overline{X} - \mu}{\sigma / \sqrt{n}}$$

dmacpher@coss.fsu.edu

Jump to first page

Sampling of Means Example

- Örnek için Ortalama (X) ~N(75,25)dir.16 öğrencinin ortalamasının 78 den büyük olma olasılığı nedir?

$$\overline{X} \sim N(75, 25/16)$$

$$Z = \frac{\overline{X} - \mu}{\sigma / \sqrt{n}} = \frac{78 - 75}{5 / \sqrt{16}} = 2.4$$

$$P(\overline{X} > 78) = P(Z > 2.4)$$

$$= 0.5000 - 0.4918 = 0.0082 \quad \text{(Table A-1)}$$

dmacpher@coss.fsu.edu

Jump to first page

Merkez limit teoremi

- Rastgele değişkenlerden örnek anlam normal dağılmış PDF ten mutlaktır.
- Toplam kütle normal değildir.

dmacpher@coss.fsu.edu
Jump to first page

12. Ki kare dağılımı

dmacpher@coss.fsu.edu
Jump to first page

Ki kare dağılımı

- rastgele değişkenlerin varyanslarıyla dağılmış hipotez testi için yararlı dağılım.
- Distribution of these squared terms approximates the chi-squared distribution.

dmacpher@coss.fsu.edu

Jump to first page

Ki kare dağılımı

- Eğer Z standart normal değişkense (i.e. with μ=0 and σ=1), sonra Z^2 k serbestlik derecesiyle ki kare dağılımına sahiptir.

$$\Sigma_{ʒ} = X_{ʒ}^{\kappa}$$

Serbestlik derecesi(k) hesapladığımız gözlem miktarını bağımsızlık numaralarını açıklar

dmacpher@coss.fsu.edu

Jump to first page

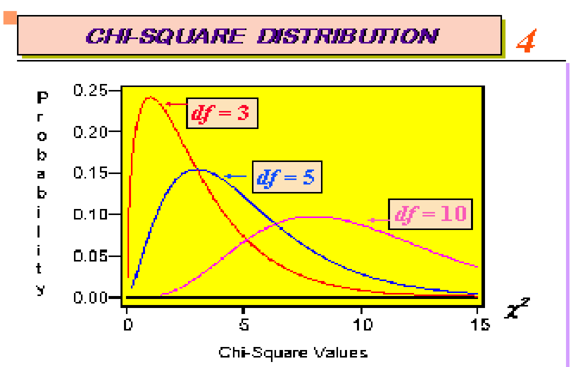
Ki kare dağılımı

- Ki kare orjinden başlar.
- Dağılımın şekli k ya bağlıdır.
  - Artan k dağılımı daha simetrik yapar..
  - k geniş oluncaki kare hemen hemen normaldir.
- Dağılımın anlamı k dır, ve varyansı 2k dır.

dmacpher@coss.fsu.edu

Jump to first page

Ki kare dağilımı



dmacpher@coss.fsu.edu

Jump to first page

Ki kare dağılımı

- Tablo A-3

k\Q	0.250	0.100	0.050	0.025	0.010	0.005	0.001
1	1.323	2.706	3.841	5.024	6.635	7.879	10.828
5	6.626	9.236	11.071	12.833	15.086	16.750	20.515
10	12.549	15.987	18.307	20.483	23.209	25.188	29.588
20	23.828	28.412	31.410	34.170	37.566	39.997	45.315
30	34.800	40.256	43.773	46.979	50.892	53.672	59.703
40	45.616	51.805	55.758	59.342	63.691	66.766	73.402

verilen 20 serbestlik derecesinde X2 karenin değeri 40 ve daha üzeri olma olasılığı ne?

As seen from table A-3, probability is 0.005 - rather small.

dmacpher@coss.fsu.edu

Jump to first page

Örnek vs doğru varyans

- To find the probability of obtaining a given sample σ^2 compared to the true σ^2:
  - Varyansla normal toplamı farz edelim: σ^2
  - Bu toplamdan varyansla rastgele örnek almak S^2.

$$\text{Then } (n-1)\left(\frac{S^2}{\sigma^2}\right) \sim X^2_{(n-1)}$$

dmacpher@coss.fsu.edu

Jump to first page

Örnek vs doğru varyans

- 25 sınav sonucunu rastgele aldığımızı farzedelim ve varyans 25 tir.örnek değerin olasılığı nedir?Suppose.
- Doğru varyans 20 dir. What is the probability of obtaining such a sample value?
  - (24) (25/20) = 30
  - The probability of obtaining a X^2 of 30 for 24 d.f. is between 10% and 25%.

dmacpher@coss.fsu.edu

Jump to first page

13. t-dağılımı

dmacpher@coss.fsu.edu

Jump to first page

t dağılımı

- varyans bilinmezken hipotezi test etmek için t dağılımını kullanırız.

Remember that if $\overline{X} \sim N(\mu, \sigma^2/n)$, then

$$Z = \frac{\overline{X} - \mu}{\sigma/\sqrt{n}} \sim N(0,1)$$

Suppose only know μ and don't know σ^2,

we only know its estimator S^2.

dmacpher@coss.fsu.edu

Jump to first page

t dağılımı

- Remember $S^2 = \frac{\sum (X_i - \overline{X})^2}{(n-1)}$, so define a new variable

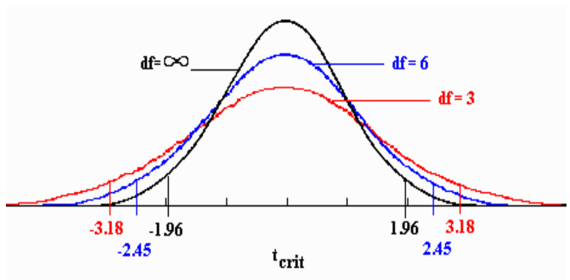
$$t = \frac{\overline{X} - \mu}{S/\sqrt{n}}$$

T dasğılımının devamı için kitabın arkasındaki tablo A- 2 ye bakın.

dmacpher@coss.fsu.edu

Jump to first page

t dağılımı

- 

dmacpher@coss.fsu.edu

Jump to first page

t dağılımının özellikleri

- Normal gibi, t dağılımı simetriktir ve geniş örnekler içinnormale yaklasıktır.
  - Örnek kütle 30dan küçükken daha geniştir.
  - 30 üzerindeki örnek kütlede,t normale yaklasıktır..

dmacpher@coss.fsu.edu

Jump to first page

t Dagılımının özellikleri

- t dağılımının anlamıstandart normal dağılımdaki gibi sıfırdır.
- varyansı k/(k-2).
  - Geniş örnek kütleler için, the varyans standart normal dağılımdaki gibi 1 e yaklasır.

dmacpher@coss.fsu.edu

Jump to first page

t dağılım tablosu

- Table A-2

k\Q	0.500	0.200	0.100	0.050	0.025	0.010	0.002
1	1.000	3.078	6.314	12.706	31.821	63.657	318.310
5	0.727	1.476	2.015	2.571	3.365	4.032	5.893
10	0.700	1.372	1.812	2.228	2.764	3.169	4.144
20	0.687	1.725	2.086	2.131	2.528	2.845	3.552
30	0.683	1.310	1.697	2.042	2.457	2.750	3.385
Infinity	0.674	1.282	1.645	1.960	2.326	2.576	3.090

Given df=20, what is the probability of obtaining a t value of 2.85 or greater = 0.01

dmacpher@coss.fsu.edu

Jump to first page

t dağılım tablosu

- %5 anlamlılık testi için,t dağılımının kritik değeri 1.96 ya yaklasır.normal dağılımın krıtık değeri n geniş olunca olur.
- 30 veya üzeri örnek için, 2 nin kritik değeri akla uygundur.

dmacpher@coss.fsu.edu

Jump to first page

t dağılımı örneği

- 25 sınav sonucunun rastgele değişkenlerine sahibiz anlam:78 ve S^2 :16.N 75 anlamlılığında örnek ortalama olasılığı nedir?

$$t = \frac{\overline{X} - \mu}{S/\sqrt{n}} = \frac{78-75}{4/\sqrt{25}} = \frac{3}{4/5} = 3.75$$

Probability of obtaining a t-value this big or greater for 24 df is much less than 1%.

dmacpher@coss.fsu.edu

Jump to first page

14. F dağılımı

dmacpher@coss.fsu.edu

Jump to first page

F Distribution

- X ve Y toplam kütleden iki bağımsız rastgele değişken sahiibiz
- $X \sim N(\mu_X, \sigma^2_X)$ and $Y \sim N(\mu_Y, \sigma^2_Y)$
- Varyanslarının aynı olup olmadığını test etmek istiyoruz..
  - Tahmın edilmiş varyanslar S^2_X ve S^2_Y
  - oran $S^2_X / S^2_Y \sim F$ dır her iki ana kütle eşitse.

dmacpher@coss.fsu.edu

Jump to first page

F Distribution

- F-dağılımını $F_{m,n}$ ile belirtiyoruz.serbestlik derecesi iki değişkenle tanımlandığı zaman, m numerator oluyor ve n serbestlik derecesidir
- F daima büyüktür1den.
  - If varyanslar eşitse, F 1 düzeyündedir.
  - F nin daha geniş değeri, varyanslar eşittir.

dmacpher@coss.fsu.edu

Jump to first page

F Dağılımı

- F dağılımı is çarpıktır, ki karede olduğu gibi.
- Sıralaması sıfır ve sonsuz arasında.
- m ve n büyürken,f testi normale yaklasır.

dmacpher@coss.fsu.edu

Jump to first page

5% F dağılım tablosu A-4

n\m	1	2	3	4	5	6	7
1	161.40	199.50	215.70	224.60	230.20	234.00	236.80
5	6.61	5.79	5.41	5.19	5.05	4.95	4.88
10	4.96	4.10	3.71	3.48	3.33	3.22	3.14
20	4.35	3.49	3.10	2.87	2.71	2.60	2.51
30	4.17	3.32	2.92	2.69	2.53	2.42	2.33
Infinity	3.84	3.00	2.60	2.37	2.21	2.10	2.01

Given m = 7 and n = 10, what is the probability of obtaining an F value of 3.14 or greater = 0.05

dmacpher@coss.fsu.edu

Jump to first page

F dağılım örneği

- Suppose test whether the variance of the SAT math test differs from the variance of the SAT verbal test.
- 21 öğrencinin örneği kullanılıyor, S^2 matematik testinde buluruz ve sözlü 100,000 ve 80,000.
- Normal dağılan test kütlesi olduğunu farz edelim, f istatistiği 1.25 (100,000/80,000) ile20 ve20 serbestlik derecesidir..

dmacpher@coss.fsu.edu

Jump to first page

F dağılım örneği

- Elde edilen 1.25 in Fdeğeri serbestlik dercesinden fazladır
- İki varyansın eşit olduğu kuralı uygulayamayız
  - If we had come up with a higher number - e.g. 3.0 the probability of obtaining this is less than .01
    - Ana kütle varyasının eşit olduğunu söyleyemeyız.

dmacpher@coss.fsu.edu

Jump to first page

15. İstatiksel sonuç

dmacpher@coss.fsu.edu

Jump to first page

İyi tahmin nedir?

- kriter:
  - Computational Cost
  - yansız
  - Hızlı ve verimli
  - Best Linear Unbiased Estimator
  - tutarlı

dmacpher@coss.fsu.edu

Jump to first page

Computational Cost

- Some methods involve a lot more time and effort than others.
- More advanced techniques can often considerable extra effort - is it worth it?
- Fortunately, many types of models are now routinely available on most statistical packages, so this criteria is less important than it used to be.

dmacpher@coss.fsu.edu

Jump to first page

yansız

- 1000 örnek aldığımızı farz edelim
- Her örnek için örnek anlamın tahmin edicisine sahip olmamız gerekir.
- Bu tahminler aynı almayacaktır, normale yaklasan sağılımın örneğine yaklasacaktır.

dmacpher@coss.fsu.edu

Jump to first page

yansız

- So in repeated sampling the mean value of all these estimators tends towards the true population mean, μ
  - Örnek anlam yansız tahmin edici olarak tanımlanır.çünkü:

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

dmacpher@coss.fsu.edu

Jump to first page

Verimli

- Birden fazla yansız tahmınedene sahip olabiliriz – arasından hangini seçebiliriz?
  - Enküçük varyanslı örnek kütleninki seçilir..
  - Geniş varyansla tahmin ediciden tahmın yapmak daha az istenir,

dmacpher@coss.fsu.edu

Jump to first page

verimli

- Tahmin edici sıfır varyansa sahipse,parametreleri tam bulabiliriz.
- Yansız tahmin ediciyi enküçük varyansla seçilir. En yansız veya en verrimli tahmin edicidir.

dmacpher@coss.fsu.edu

Jump to first page

En doğrusal yansız tahmin edici

- Bu ekleme tahmin edicinin doğrrusal olmasını sınırlar.
- örneğin, örnek anlam doğrusal tahmin edicidir.

$$\overline{X} = \sum_{i=1}^{n} X_i / n$$

- This makes it easier to figure out which is the most efficient estimator

dmacpher@coss.fsu.edu

Jump to first page

En doğrru yansız tahmin edici

- An estimator that is linear, unbiased, and has minimum variance among all linear unbiased estimators is called the best linear unbiased estimator - BLUE.
- Parametrelerin teknumaralı olanları ile ilgileniyouz.
  - It is harder to find the BLUE in the the multidimensional case.

dmacpher@coss.fsu.edu

Jump to first page

tutarlı

- Örnek kütle genişl olduğunda tahmin edici ne olur?
  - Tahmin edici tutarlı ise örnek kütle daha geniş olur, toplam parametrelerin doğrru değerine yaklaşır..

$\hat{\beta}$ is a consistent estimator of β if $\text{plim}\ \hat{\beta} = \beta$

where plim is the probabilit y limit

dmacpher@coss.fsu.edu

Jump to first page

Consistent

- Sometimes we can get consistency but not unbiasedness, where the estimator may not equal the true parameter on average, but will approximate the true parameter as sample size gets large.

dmacpher@coss.fsu.edu

Jump to first page

16. Estimation

dmacpher@coss.fsu.edu

Jump to first page

Estimators

- An estimator is generated when we take a random sample and compute, for example, the mean or variance.
- Typically, we do not know the mean and variance of the population, so we have to rely on our sample estimates.

dmacpher@coss.fsu.edu

Jump to first page

Point Estimates

- Suppose we take a random sample of 16 students and get an average of 75 and variance 25.
  - The sample mean of 75 is the point estimate of μ the population parameter,which is typically unknown.
- This point estimate will vary across samples, since the sample mean is an r.v.

dmacpher@coss.fsu.edu

Jump to first page

Interval Estimates

- An alternative is that a certain interval contains the true mean.

Remember that if $\overline{X} \sim N(\mu, \sigma^2/n)$, then

$$Z = \frac{\overline{X} - \mu}{\sigma/\sqrt{n}} \sim N(0,1)$$

Suppose don't know σ^2,

we only know its estimator S^2.

dmacpher@coss.fsu.edu

Jump to first page

Interval Estimates

If we use S^2, then we know $t = \dfrac{\overline{X} - \mu}{S / \sqrt{n}}$

which follows the t distribution with n - 1 d.f.

dmacpher@coss.fsu.edu

Jump to first page

Interval Estimates Example

- Suppose we want to construct a 95% confidence interval around this mean value of 75
  - There is a 95% probability that this interval will contain the true μ.
- Check t-tables for a 5% 2-tailed test.
  - For 15 df - then P(|t|>2.131) = .05
    - So P(-2.131<t<2.131) = .95

dmacpher@coss.fsu.edu

Jump to first page

Interval Estimates Example

- Substitute in our definition of t:

$$P(-2.131 < \frac{\overline{X} - \mu}{S / \sqrt{n}} < 2.131) = .95$$

$$P(\overline{X} - \frac{2.131S}{\sqrt{n}} < \mu < \frac{2.131S}{\sqrt{n}}) = .95$$

Above is a confidence interval - there is a 95% probability that this interval contains the true population mean.

dmacpher@coss.fsu.edu

Jump to first page

Interval Estimates Example

- Plugging in the values, the confidence interval is:

$$P(\overline{X} - \frac{2.131(5)}{\sqrt{16}} < \mu < \overline{X} + \frac{2.131(5)}{\sqrt{16}}) = .95$$

$$(75 - 2.66) < \mu < (75 + 2.66)$$

72.34<μ<77.66 is the 95% confidence interval for μ

So there is a 95% probability that this interval contains the true mean

dmacpher@coss.fsu.edu

Jump to first page

Alpha

- Sometimes the confidence interval is expressed as:
  - P(lower limit< μ<upper limit) =1-α
    - 1- α is the probability that the random interval contains the true μ (the confidence coefficient).
- α is known as the level of significance
  - Here α is .05 or 5% (the probability of committing a Type I error).

dmacpher@coss.fsu.edu

Jump to first page

17. Hypothesis Testing

dmacpher@coss.fsu.edu

Jump to first page

Hypothesis Testing

- We can evaluate the relationship between the sample estimates and the population parameters.
- Suppose a random sample has a mean of 75 on an exam.
  - Hypothesize the true mean(μ)=78.
    - Is 75 statistically different from 78?
    - Is the 75 due to sampling error?
    - Is the 75 because μ is not 78?
    - This is hypothesis testing.

dmacpher@coss.fsu.edu

Jump to first page

Hypotheses

- Set up a null hypothesis.
  - Ho: μ = 78
- We test this against the alternative hypothesis H1.
  - H1: μ is not equal to 78
- Alternative hypotheses for H1:
  - H1: μ > 78
  - H1: μ < 78
    - These are one-sided hypotheses

dmacpher@coss.fsu.edu

Jump to first page

Confidence Interval Approach

- We have constructed a confidence interval around μ
  - The 95% confidence interval was: $72.34 < \mu < 77.66$
  - The confidence interval indicates there is a 95% probability that this interval contains the true mean.
  - Our null hypothesis is that $\mu = 78$
  - It does not lie within the interval.

dmacpher@coss.fsu.edu

Jump to first page

Confidence Interval Approach

- We can reject the hypothesis that μ is 78.
  - The 95% interval is the acceptance region.
  - The area outside it is the critical region, or rejection region.
    - 78 is inside the critical region - the probability of μ being 78 is less than 2.5%

dmacpher@coss.fsu.edu

Jump to first page

Confidence Interval Approach

- We reject the null that μ is 78.
  - We are only 95% confident in this
  - Perhaps our sample did come from a population with $\mu = 78$.
  - If so we have committed a type I error
    - We have rejected a hypothesis that is actually true.

dmacpher@coss.fsu.edu

Jump to first page

Confidence Interval Approach

- Suppose we want to reduce the risk of committing a Type I error.
- Suppose we want to fix α at 1%, not 5%.
  - We construct a 99% confidence interval around μ
- Use t tables for 1% two-tailed test
  - For 15 df - then $P(|t|>2.947) = .01$
  - So $P(-2.947<t<2.947) = .01$

dmacpher@coss.fsu.edu

Jump to first page

Confidence Interval Approach

- Substitute in our definition of t:

$$P\,(-2.947 < \frac{\overline{X} - \mu}{S / \sqrt{n}} < 2.947) = .99$$

$$P\,(\overline{X} - \frac{2.947S}{\sqrt{n}} < \mu < \frac{2.947S}{\sqrt{n}}) = .99$$

dmacpher@coss.fsu.edu

Jump to first page

Confidence Interval Approach

- Plugging in the values, the confidence interval is:

$$P\,(\overline{X} - \frac{2.947(5)}{\sqrt{16}} < \mu < \frac{2.947(5)}{\sqrt{16}}) = .99$$

$$(75 - 3.68) < \mu < (75 + 3.68)$$

71.32<μ<78.68 is the 99% confidence interval for μ

Our hypothesis is not rejected - because we have minimized the probability of committing a Type I error.

dmacpher@coss.fsu.edu

Jump to first page

Type 1 and Type 2 Errors

- Type I error = α (the level of significance)
  - Reject Ho when it is true
  - 5% level of significance - the same thing as a 95% level of confidence.
- Type II error = β (accept Ho when it is false).
- The confidence coefficient is 1- α
- The power of the test is 1- β

dmacpher@coss.fsu.edu

Jump to first page

Type 1 and Type 2 Errors

- Switch of α from 5% to 1%, increased the probability of committing a Type II error
  - Accepting a false hypothesis.
- In practice we generally try to minimize type I errors
  - Keep the probability of committing such an error at a low level such as 1% or 5%

dmacpher@coss.fsu.edu

Jump to first page

Test of Significance Approach

- An alternative, complementary approach.

We know $t = \dfrac{\overline{X} - \mu}{S/\sqrt{n}}$,

which follows the t distribution with n - 1 d.f.

- Suppose we take our random sample of 16 students, mean of 75 and variance 25.

dmacpher@coss.fsu.edu

Jump to first page

Test of Significance Approach

- So, Xbar= 75, S=5, n=25
- We don't know μ, so specify a value for μ.
- Set up hypothesis test:
  - H0: $\mu = 78$
  - H1: $\mu \neq 78$
- Calculate t:
  - t = (75-78)/(5/4) = -2.400

dmacpher@coss.fsu.edu

Jump to first page

Test of Significance Approach

- For α= 5% and df of 15, the critical 2 tailed t values are + or -2.131.
  - The probability of getting a t higher than 2.131 or lower than -2.131 is only 5%
- Our t-value is lower than this, so we can reject the null hypothesis that the true μ is 78 at the 5% level.

dmacpher@coss.fsu.edu

Jump to first page

Test of Significance Approach

- Suppose we choose an α of 1%; the critical t values are + or - 2.947.
- Our t-value is within this range, so we cannot reject the null at the 1% level of significance.

dmacpher@coss.fsu.edu

Jump to first page

t-test

- This test is called the t-test.
  - We will use it as a test of significance of slope parameters in regression.
- Statistically significant means we can reject the null hypothesis.
- Commonly, we choose 1%, 5% and 10% levels of significance.

dmacpher@coss.fsu.edu

Jump to first page

One-tailed test

- Can also do a one-tailed test.
  - H0: $\mu = 78$
  - H1: $\mu < 78$
- Again, we choose 5% level of significance and find the t-value is <-1.753
- So again reject the H0 that the mean equals or is greater than 78.

dmacpher@coss.fsu.edu

Jump to first page

X^2 test of significance

- A random sample of 10 SAT scores has a variance of 84.
- Test the hypothesis that the true variance is 86.
- Ho: $\sigma^2 = 86$
- H1: $\sigma^2 \neq 86$

Remember $(n-1)\left(\frac{S^2}{\sigma^2}\right) \sim X^2_{(n-1)}$

dmacpher@coss.fsu.edu

Jump to first page

X^2 test of significance

- Choose 5% significance level
  - X^2 = 9(84/86) = 8.79
- For df = 9 and 5% level of significance, we would need a critical value of X^2 =16.92
  - Our X^2 is much less than that, so we cannot reject the null.
- An alternative: the probability of obtaining X^2 of 8.79 is about 50%

dmacpher@coss.fsu.edu

Jump to first page

F test of significance

- Remember $S^2_X / S^2_Y \sim F_{m-1,\, n-1}$ if the variances of the two populations are equal.
- Assume two random samples of students in two different schools taking a test:
  - $S^2_X = 9.0$ with sample size 51
  - $S^2_Y = 7.2$ with sample size 41
- Ho: $\sigma^2_X = \sigma^2_Y$

dmacpher@coss.fsu.edu

Jump to first page

F test of significance

- F = 9/7.2 = 1.25
  - df numerator 50
  - df denominator 40
- 5% level critical F value is 1.66
  - Since our actual F is less than that, we cannot reject the null that the two population variances are the same.

dmacpher@coss.fsu.edu

Jump to first page